Pricing the Left Tail: Consumption Skewness and Expected Returns *

C. Christopher Hyland† Niklas Schmitz‡

November 1, 2026

Abstract

Consumption-based asset pricing implies that risks to future consumption growth should be priced. We estimate the conditional distribution of future U.S. consumption growth by emulating investors' forecasting problem with quantile regressions and machine learning on a large macroeconomic panel, and construct several risk measures from it. Asymmetry, measured by Kelley skewness, predicts equity returns: a one-standard-deviation fall in Kelley skewness forecasts 5 percentage points higher 12-month excess returns. Variance and dispersion have no comparable predictive power. Kelley skewness betas explain 91% of the variation in average returns across size, book-to-market, and industry portfolios. Among macroeconomic shocks, financial uncertainty and adverse investment-specific technology news lower Kelley skewness the most, mainly through the lower tail. Monetary easing is associated with lower downside risk when that risk is already high, such as in recessions, but not when it is low.

Keywords: Asset Pricing, Macro-Finance, Higher-Order Moments.

*We are grateful to Yacine Aït-Sahalia, Harjoat Bhamra, Rhys Bidder, Giancarlo Corsetti, Radu Cristea, Elisa Faraglia, Petra Geraats, Chryssi Giannitsarou, Tobias Kawalac, Zach Mazlish, Alexander Meyer-Gohde, Daniel Ostry, Charles Parry, Alba Patozi, Carolin Pflueger, and Giovanni Ricco for helpful comments. We thank Jakob Berndt, Refet S. Gürkaynak, Marek Jarociński, and Ian Martin for sharing data. All errors are ours.

†University of Oxford, Saïd Business School, Exeter College, and Oxford-Man Institute. Email: Christopher.Hyland.DPHIL@said.oxford.edu.

‡Point72. The information, views, and opinions expressed herein are solely Niklas Schmitz's and do not necessarily represent the views of Point72 or its affiliates. Point72 and its affiliates are not responsible for, and did not verify for accuracy, any of the information contained herein.

1 Introduction

The standard consumption-based asset pricing model (Lucas, 1978; Breeden, 1979) implies that the conditional expectation of future excess returns can be approximated as a function of risks to future consumption growth. Let equity returns between period t and $t+1$ be $r_{t,t+1}$ such that expected excess returns over the risk-free rate r^f are $E_t\left(r_{t,t+1}\right) - r^f_{t,t+1}$. Given a stochastic discount factor $m_{t,t+1}$ and consumption growth rate $c_{t,t+1}$, expected excess returns can be approximated as (Cochrane, 2009):

$$E_t\left(r_{t,t+1}\right) - r^f_{t,t+1} = -\frac{\mathrm{cov}_t\left(m_{t,t+1}, r_{t,t+1}\right)}{E_t\left(m_{t,t+1}\right)} \approx \gamma_t \sigma_t\left(c_{t,t+1}\right) \sigma_t\left(r_{t,t+1}\right) \rho_t\left(c_{t,t+1}, r_{t,t+1}\right), \tag{1}$$

where γ_t is risk aversion and the last term is the conditional correlation between consumption growth and returns. The term $\sigma_t\left(c_{t,t+1}\right)$ captures expected risks to consumption growth between t and $t+1$. According to the theory, an increase in expected risks to future consumption growth raises expected excess returns, because investors demand a higher equity premium.

Since equation (1) holds in expectation conditional on information available at time t, testing it requires a measure of conditional consumption risk. The measure should reflect investors' assessment of the distribution of possible future consumption growth, given current economic conditions. Such a measure is not readily available. Survey evidence on household consumption risk expectations has only recently become available and covers short samples (Christelis et al., 2020; Guiso and Jappelli, 2024). While the theory predicts that market prices themselves contain the relevant measure of consumption risk, we cannot extract this information without taking a stand on functional forms and parameter values, which are still actively debated (Cochrane, 2017). This leaves a gap in empirical tests of consumption-based asset pricing models, which have largely relied on realized consumption growth and therefore mix its expected and unexpected components.

This paper estimates the conditional distribution of consumption growth, constructs risk measures based on it, and studies their asset pricing performance. First, we emulate the out-of-sample forecasting problem that investors face. Each month since 1970, we use quantile regressions to forecast different scenarios for year-ahead U.S. aggregate consumption growth, using only data up to that month, and approximate the full conditional distribution from them. Second, we use the conditional distribution to construct a variety of risk measures, such as the mean, the variance, and measures of asymmetry. Third, we test whether these measures predict excess equity returns and price the cross section of returns. The main finding is that skewness in the conditional distribution, measured by

Kelley skewness, predicts excess returns in the time series and is priced in the cross section.[1] Low values of Kelley skewness indicate large downside risks to future consumption growth and predict higher excess returns, in line with the theory.

Estimating the conditional distribution of consumption growth is complicated because it can depend on many state variables, each of which may have a small but non-negligible effect (Giannone et al., 2021). Investors face the same high-dimensional forecasting problem (Martin and Nagel, 2022). To address it, we use methods from the machine learning literature and, similar to Bianchi et al. (2022), interpret the machine learning forecast as a benchmark that a rational investor would want to achieve given the same information. We forecast the quantiles out of sample in two steps. A pre-selection step uses ℓ_1-penalized quantile regression to narrow down the 126 series of the FRED-MD panel (McCracken and Ng, 2016). A targeted factor estimation step, the partial quantile regression of Giglio et al. (2016), then aggregates their information, weighting each selected predictor by its forecasting power for the quantile being targeted. To recover the full conditional distribution from the quantiles, we fit a flexible skewed t-distribution across them, as in Adrian et al. (2019). The machine learning forecast is more accurate on average than simpler alternatives, with the largest gains in the center and upper part of the distribution. Using only the National Financial Conditions Index (NFCI) as a predictor, as in Adrian et al. (2019), yields an out-of-sample R^1 (the quantile analogue of R^2) of 2.6% over 1981–2025, against 5.6% for the machine learning forecast over the same sample and 11.3% over the full 1970–2025 window. Because the predicted quantiles do not achieve exact coverage, we use the fitted distribution as a time-varying description of relative tail shape rather than as a set of exactly calibrated probabilities.

The resulting estimates of the conditional distribution have several attractive features. The estimation approach is flexible and does not rely on assumptions about the parameters or functional forms of a consumption-based asset pricing model, yet it yields an estimate of the conditional distribution that maps directly into the theory. The conditional distribution varies with the business cycle. In recessions its mean falls and its variance rises, particularly in the Great Recession and the COVID-19 shock. The distribution narrowed during the Great Moderation and has been wider since the Great Recession. Kelley skewness falls in recessions, because the upper quantiles are more stable than the lower quantiles.

The approximation in equation (1) summarizes consumption risk by its standard de-

[1] Kelley skewness compares the distance from the median to the 90th percentile with the distance from the median to the 10th percentile, relative to the distance between the 90th and 10th percentiles. It is bounded between −1 and 1. A value of zero indicates a symmetric distribution. Negative values indicate a left-skewed conditional distribution, which represents large downside risk.

viation, treating good and bad outcomes symmetrically. Models of disaster risk instead suggest that downside risks to consumption growth may be the relevant driver of equity premia (Rietz, 1988; Barro, 2006; Gabaix, 2012; Wachter, 2013). To stay as agnostic as possible about the risk measure, we construct several, some from the predicted quantiles, such as dispersion and Kelley skewness, and others from the fitted distribution, such as the mean, variance, and relative entropy. The risk measures behave intuitively over the business cycle: variance and dispersion are high when economic policy uncertainty (Baker et al., 2016) and financial uncertainty (Ludvigson et al., 2021) are high, while Kelley skewness tends to be low when policy uncertainty is high.

Across the predictive return regressions, Kelley skewness emerges as the most consistent predictor. Over 1970–2025, a one-unit decrease in Kelley skewness, such as from +0.5 to −0.5, predicts 12-month excess market returns roughly 16 percentage points higher; scaled by the in-sample standard deviation of Kelley skewness, the effect is about 5 percentage points. Kelley skewness explains 9.5% of the in-sample variation in excess returns at the 1-year horizon and 16.3% at the 3-year horizon. Unlike the conditional mean, it is significant at all three forecast horizons and in every sample window we consider. It also remains significant in ten of twelve bivariate regressions, including those that control for financial uncertainty, volatility, the variance risk premium, and credit spreads. Kelley skewness predicts excess returns out of sample as well, with an R^2 of 7.8% at the 12-month horizon. This is the largest among the consumption-based predictors and above the values typically reported for valuation ratios and tail-risk indices (Campbell and Thompson, 2008; Welch and Goyal, 2008; Kelly and Jiang, 2014), although the gain over a constant-only benchmark is not statistically significant. Symmetric risk measures such as variance and dispersion, by contrast, have small and insignificant coefficients at every horizon. Investors therefore appear to be compensated for downside consumption risk, not for symmetric uncertainty about consumption.

Kelley skewness is also priced in the cross section of returns. Fama and MacBeth (1973) regressions on portfolios sorted on size and book-to-market and on industry show that assets with greater exposure to left-skewed consumption states earn higher average returns. The slope on the Kelley skewness beta is −0.51, and exposure to Kelley skewness explains 90.6% of the cross-sectional variation in average returns. The slope remains negative in every calendar-month subsample of non-overlapping annual returns and is significant at the 5% level in six of the twelve, which alleviates concerns that overlapping observations drive the result. The one-factor model does not price every portfolio exactly, however, which suggests that other risk factors are needed to span the full cross section of returns.

The conditional distribution of consumption growth is an equilibrium object that is ultimately determined by the shocks hitting the economy. To shed some light on its drivers, we study which shocks move Kelley skewness. In Section 4, we estimate local projections for four candidates from the macro-finance literature: tighter credit conditions (Gilchrist and Zakrajšek, 2012; Brunnermeier and Sannikov, 2014; Adrian et al., 2022), higher financial uncertainty (Bloom, 2009; Jurado et al., 2015; Bloom et al., 2018; Ludvigson et al., 2021), adverse oil-supply news (Hamilton, 2003; Kilian, 2009; Känzig, 2021), and adverse news about investment-specific technology (IST), which lowers expected productive capacity (Beaudry and Portier, 2006; Ben Zeev and Khan, 2015). Financial uncertainty moves Kelley skewness the most. A one-standard-deviation increase lowers it by about 9 percent of its standard deviation over the following two to six quarters. Adverse IST news and tighter credit conditions lower Kelley skewness by less, and oil-supply news barely moves it. The shocks also differ in how they move the distribution. Financial uncertainty and adverse IST news lower the 10th percentile of expected consumption growth by more than the median, so they stretch the lower tail. Credit tightening instead shifts the whole distribution down.

We study monetary policy separately, using high-frequency surprises (Gürkaynak et al., 2005; Gertler and Karadi, 2015; Nakamura and Steinsson, 2018; Jarociński and Karadi, 2020; Miranda-Agrippino and Ricco, 2021). *A priori*, its relationship with downside consumption risk is ambiguous. In crisis times, a key goal of monetary policy is to mitigate downside risks and avert bad outcomes, and accommodative policy can help by easing financial constraints and supporting demand (Borio and Zhu, 2012; Bekaert et al., 2013). But excessively loose financial conditions can also contribute to the buildup of leverage and leave the economy more vulnerable to later shocks (Schularick and Taylor, 2012; Brunnermeier and Sannikov, 2014; Adrian et al., 2022). Loose monetary policy may therefore either decrease or increase downside risk, depending on the state of the economy.

On average, monetary easing barely moves Kelley skewness. The response depends, however, on how high downside risk is when the surprise occurs. When recent Kelley skewness is low, a one-standard-deviation easing surprise raises it by about 0.05 over the following six to eighteen months. When recent Kelley skewness is high, the same surprise lowers it by about 0.02. This pattern holds for nearly all of the eleven high-frequency surprise series we consider, and across sample windows and definitions of the state (Appendix E). Most of the improvement occurs in the lower tail: when downside risk is high, easing raises the 10th percentile of expected consumption growth by more than three times as much as the median. These estimates are associations and cannot tell

the two channels apart. But they fit both views. Easing is associated with lower downside risk when that risk is already high, which is also when expected equity returns are highest, and with higher downside risk when it is low.

Related literature. Our work contributes to several strands of literature. The focus on risks to aggregate consumption is rooted in consumption-based asset pricing (Lucas, 1978; Breeden, 1979); see Cochrane (2017) for a review. Many empirical tests of the consumption-based model choose a measure of realized consumption growth that rationalizes the theory in the data (Mankiw and Shapiro, 1986; Breeden et al., 1989; Parker and Julliard, 2005; Savov, 2011; Kroencke, 2017). Two exceptions focus on conditional moments of consumption growth. Kandel and Stambaugh (1990) estimate the conditional mean and variance of consumption growth in sample and find only weak evidence that they forecast returns. Boguth and Kuehn (2013) estimate a Markov chain for the first two moments of consumption growth and find that consumption volatility is priced in the cross section. We contribute to this literature by estimating the full conditional distribution of consumption growth out of sample from a large set of predictors. Its Kelley skewness predicts returns even after controlling for the conditional mean, while its variance has no predictive power.

The predictive power of asymmetry in the conditional distribution is connected to, but distinct from, the disaster risk literature, which explains the equity premium by the risk of rare but very large recessions (Rietz, 1988; Barro, 2006). Gabaix (2012), Gourio (2012), Nakamura et al. (2013), and Wachter (2013) extend this idea in different frameworks, including time-varying disaster risk. Relative to this literature, we find evidence for the importance of downside as opposed to symmetric risks in the growth outlook. These risks vary over ordinary business cycles and do not rely on extreme outcomes, because Kelley skewness is measured at the 10th and 90th percentiles. This is in line with Beason and Schreindorfer (2022), who find that monthly market returns below -10 percent account for about 80 percent of the equity premium, while returns below -30 percent matter little. Investors are therefore compensated mainly for large – but not extreme – declines in equity prices. Measuring downside risk at these percentiles also avoids the difficulty of estimating extreme tails from short samples of macroeconomic data.

Our results also relate to work on return prediction with risk measures derived from financial-market information. Option-based measures such as the variance risk premium (Bollerslev et al., 2009) and tail risk (Bollerslev and Todorov, 2011b; Bollerslev et al., 2015) predict equity premia, and Kelly and Jiang (2014) estimate tail risk from the cross section of individual returns and find that it predicts market returns. Because they are inferred from asset prices, these measures combine expectations about fundamentals

with compensation for bearing risk (Bekaert and Hoerova, 2014). Kelley skewness is estimated from macroeconomic data instead, so it measures the quantity of consumption risk, whether that risk originates in financial markets or in the real economy. It predicts returns even after controlling for the variance risk premium and the VIX.

Our estimation approach builds on work in macroeconomics that forecasts quantiles of the distribution of a variable of interest and constructs risk measures from them. Giglio et al. (2016) use partial quantile regression to study how well systemic risk indices forecast negative shocks to macroeconomic outcomes. Adrian et al. (2019) estimate the conditional distribution of GDP growth with quantile regressions and fit a skewed t-distribution across the estimated quantiles. Because they focus on the effect of financial conditions on downside risk, they use the NFCI as the sole predictor. We forecast consumption growth, because consumption is the marginal-utility object in the asset pricing model, and draw on the full FRED-MD panel. The difference matters for returns. An NFCI-only forecast captures downside financial-stress states well but describes the center and upper tail less accurately, which can overstate the asymmetry of the outlook, and its Kelley skewness does not predict returns. Consumption Kelley skewness also retains its predictive content after controlling for the New York Fed's Outlook-at-Risk measure of GDP growth at risk, while that benchmark does not (Appendix D). More broadly, conditional macroeconomic distributions carry asset pricing information beyond their means (Bauer and Chernov, 2024; Hyland, 2026).

The monetary-policy results contribute to a growing literature on monetary policy and downside risks to growth. Adrian et al. (2019) find that financial conditions predict downside risk to GDP growth, and Loria et al. (2023) estimate stronger effects of monetary shocks on the lower than on the upper quantiles of the growth distribution. Duprey and Ueberfeldt (2020), Adrian et al. (2022), and Ajello and Pike (2022) stress an intertemporal trade-off for central bankers: a loosening in financial conditions reduces growth risks in the short term but increases them in the medium term. We find that the association between easing and downside consumption risk also depends on how high that risk already is. Work on nonlinear monetary transmission has focused on other states, such as the business cycle, the sign of the shock, and its size (Tenreyro and Thwaites, 2016; Barnichon and Matthes, 2018; Ascari and Haber, 2022). A separate strand shows that monetary policy moves risk premia, risk aversion, and market-implied tail risk in financial markets (Bekaert et al., 2013; Hattori et al., 2016; Nagel and Xu, 2024; Knox et al., 2024). Our analysis differs in using a risk measure based on macroeconomic fundamentals rather than financial-market data, which matters because monetary policy also moves risk aversion (Bekaert et al., 2013).

Outline. Section 2 estimates the conditional distribution of consumption growth and constructs the candidate risk measures. Section 3 tests which of these measures predict equity returns in the time series and price the cross section. Section 4 studies which shocks move the priced measure and examines the state-dependent role of monetary policy. Section 5 concludes.

2 Estimating conditional consumption risk

Because theory does not identify the correct risk metric *a priori*, we construct several measures of expected risk to U.S. aggregate consumption growth. Each measure is based on the conditional distribution of future consumption growth. We first forecast several quantiles from consumption growth and a set of predictors, yielding quantile-based risk measures. We then fit a flexible density across the estimated quantiles, following Adrian et al. (2019), and use it to construct additional distribution-based measures.

To estimate the quantiles accurately, we enrich the quantile regressions with preselection and factor estimation. This contrasts with earlier approaches such as Adrian et al. (2019), who use the NFCI as the sole predictor. Our approach forecasts more accurately on average, and its risk estimates are more closely aligned with other measures of uncertainty. Section 3 also shows that risk measures based on the NFCI-only forecast do not predict equity returns, whereas those based on our forecast do.

The asset pricing tests require the conditional distribution of future aggregate consumption growth that investors could form from time-t macroeconomic information. Our statistical procedure estimates that object directly. The resulting risk measures give the probability of a bad draw for aggregate consumption conditional on current information, which maps into the risk term in consumption-based asset pricing models. Newspaper-based uncertainty indices and most qualitative surveys measure broader perceptions of economic conditions rather than this distribution.[2] Recent surveys provide important exceptions. Christelis et al. (2020) elicit the household-specific distribution of future consumption from a representative Dutch household panel, and Guiso and Jappelli (2024) elicit probability distributions for expected household consumption growth in the Italian Survey of Consumer Expectations. These studies recover theory-relevant moments of household consumption expectations, but each contains only three survey waves. Our monthly index begins in 1970 and therefore covers several U.S. business cycles and recessions. It also estimates the conditional physical distribution of aggregate consumption

[2]For example, a typical question in the Michigan Consumer Survey is "Would you say that at the present time business conditions are better or worse than they were a year ago?"

growth from macroeconomic data. Survey answers about future household consumption, which is partly a choice variable, may combine physical risk expectations with risk aversion.

Because our measures are built from the macroeconomic outlook, they can distinguish recessions from pure financial stress. Market-based risk measures, by contrast, may be affected by idiosyncratic firm effects (Kelly and Jiang, 2014) or rest on high-frequency market movements with unclear relations to macroeconomic fundamentals (Bollerslev and Todorov, 2011a). Without assumptions on the pricing kernel, they also cannot separate risk from risk aversion: measures of tail risk or risk aversion inferred from financial prices combine expectations about fundamentals with compensation for bearing risk (Bekaert et al., 2013; Hattori et al., 2016). We instead treat the estimated consumption distribution as a macroeconomic state variable that measures the quantity of consumption risk, and test its price directly in Section 3. Unlike realized consumption growth, our estimates do not use future outcomes that were unavailable when expectations were formed. Unlike symmetric measures of risk such as that of Jurado et al. (2015), they also capture asymmetries in the distribution of future outcomes. Section 3 shows that these asymmetries matter for the risk priced in equity markets.

2.1 Forecasting U.S. consumption growth: methodology

The forecasting model uses quantile regression to predict quantiles of the conditional distribution of future consumption growth. Let $c_{t,t+h}$ be aggregate real U.S. consumption growth over periods t to $t+h$ and denote the k-dimensional vector of predictors at time t by x_t. An out-of-sample (OOS) prediction of a quantile $Q(\theta)$ of the distribution of consumption growth between t and $t+h$ using data available at t is

$$\hat{Q}_{c_{t,t+h}|x_t}(\theta|x_t) = x_t{}'\hat{\beta}_{\theta,t}. \tag{2}$$

The goal is to obtain a time series of quantile forecasts for each θ_j-th quantile, where θ_j lies in a set of targeted quantile indices $\Theta = \{\theta_1, \theta_2, ..., \theta_J\}$. The remainder of this subsection describes how we obtain $\hat{\beta}_{\theta,t}$ and construct x_t.

Given consumption growth $c_{t,t+h}$ and predictors x_t, quantile regression yields the following consistent estimate for the regression coefficients:

$$\hat{\beta}_{\theta,t} = \underset{\beta_{\theta,t}}{\operatorname{argmin}} \underbrace{\sum_{s=h+1}^{t} \left(\theta \cdot \mathbb{I}_{(c_{s-h,s} \geq x'_{s-h}\beta_\theta)} |c_{s-h,s} - x'_{s-h}\beta_\theta| + (1-\theta) \cdot \mathbb{I}_{(c_{s-h,s} < x'_{s-h}\beta_\theta)} |c_{s-h,s} - x'_{s-h}\beta_\theta|\right)}_{\widehat{C}_\theta(\beta_{\theta,t})} \tag{3}$$

where $\mathbb{I}$ denotes the indicator function and $\widehat{C}_\theta$ is the sample analog of the criterion function the quantile regression estimator minimizes.

The coefficient estimates vary across quantiles $Q(\theta)$ because the objective function differs from the OLS case in two ways. First, it penalizes absolute rather than squared deviations between predicted and actual values. Second, it weights these deviations asymmetrically: observations above the fitted quantile receive weight θ and those below receive weight $1-\theta$, so that in sample roughly a share θ of observations lies below the fitted value.

The predictor vector x_t draws on the FRED-MD database of McCracken and Ng (2016). The February 2026 vintage contains 126 monthly macroeconomic series starting in January 1959. They cover output and income, the labor market, consumption and orders, inventories, money and credit, interest and exchange rates, prices, and the stock market. The dataset is updated continuously to account for changing variable definitions and data revisions.[3]

We use factor estimation to reduce the number of predictors while preserving their information content. The standard quantile regression estimator in equation (3) may not be consistent if the number of regressors is large relative to the sample size (Belloni and Chernozhukov, 2011), and a large predictor set may overfit the data. We therefore construct x_t in two steps: an algorithm first pre-selects a subset of regressors from the FRED-MD dataset, and we then construct a single factor from this subset designed to predict a consumption growth quantile of interest.

Pre-selection improves forecasting performance because some variables in the dataset are highly correlated. For example, the sample includes 13 measures of industrial production, so factors estimated from the full predictor set can place too much weight on that part of the data. Shrinking the predictor set each period also allows the factor estimates to adapt more readily to parameter changes over time (Bai and Ng, 2008). In a given period t, we pre-select a subset of predictors using ℓ_1-penalized quantile regression (Belloni and Chernozhukov, 2011). All predictor series are standardized to have zero mean and unit variance. For a given quantile index θ, the ℓ_1-penalized quantile regression coefficients $\tilde{\beta}_{\theta,t}$ satisfy

$$\tilde{\beta}_{\theta,t} = \underset{\beta_{\theta,t}}{\operatorname{argmin}} \quad \widehat{C}_\theta\left(\beta_{\theta,t}\right) + \frac{\lambda_t \sqrt{\theta(1-\theta)}}{n} \sum_{i=1}^{k} \hat{\sigma}_{i,t} |\beta_{i,t}|, \tag{4}$$

[3]The values for April 2020 were missing for the 3-month AA Financial Commercial Paper Rate (CP3Mx) and the 3-month Commercial Paper Rate minus Fed Funds rate (COMPAPFFx) due to insufficient trading data. We linearly interpolate between March and May 2020.

where $\widehat{C}_\theta$ is the sample analog of the criterion function, n is the sample size, and $\hat{\sigma}_{i,t}$ is the sample variance of predictor x_i for the sample until period t. The penalty term $\lambda_t \sqrt{\theta(1-\theta)}$ depends on the quantile index θ and a penalty level λ_t, which is re-estimated each period.[4]

ℓ_1-penalized quantile regression applies the logic of LASSO regression to a quantile regression setting. The penalty restricts the sum of absolute coefficient values, shrinking coefficients on variables with little explanatory power toward zero. For each quantile in Θ, the algorithm selects the $\tilde{k} = 40$ predictors with the largest absolute coefficients. We follow Boivin and Ng (2006), who find that factors constructed from 40 pre-screened variables can outperform factors estimated from a large dataset with over 100 macroeconomic time series.[5]

In the second step, we condense the information from the subset of regressors into a factor estimate using partial quantile regression as in Giglio et al. (2016). This approach constructs a single factor to predict the dependent variable, weighting individual predictors according to their predictive power for the dependent variable. Partial quantile regression (PQR) thus differs from principal-components factor estimation, which constructs factors that explain the most variation in the predictors. To implement the PQR, we assume that a scalar f_t contains all relevant information for a conditional quantile of consumption growth. Constructing a factor for each quantile via PQR follows two steps. First, we run univariate quantile regressions of $c_{t,t+h}$ on each predictor $x_{i,t}$ for $i = 1, ..., \tilde{k}$, with slope estimates $\hat{\gamma}_{\theta,i}$. Second, we compute the cross-sectional covariance of $x_{i,t}$ with i's first-stage slope estimate in each period t:

$$\hat{f}_{\theta,t} = \sum_{i=1}^{\tilde{k}} (x_{i,t} - \bar{x}_{i,t}) (\hat{\gamma}_{\theta,i} - \bar{\hat{\gamma}}_{\theta,i}), \tag{5}$$

where $\bar{\cdot}$ denotes the sample average. The covariance estimate $\hat{f}_{\theta,t}$ estimates the latent factor realization. It is a weighted average of individual predictors, with weights determined by their predictive strength for $Q_{c_{t,t+h}}(\theta)$ from the first step.

Given these factors, we obtain the predicted quantiles from the quantile regressions in equation (3), where the predictors are the second-stage factor estimates $\hat{f}_{\theta,t}$. We forecast the 5, 25, 50, 75, and 95% quantiles of consumption growth over the next twelve months at a monthly frequency.[6] For a given month, year-on-year consumption growth is the percent

[4] The penalty level is chosen as in Belloni and Chernozhukov (2011).

[5] More than 40 predictors typically have non-zero coefficients, so selecting 40 variables does not exclude any predictors that the ℓ_1-penalty would otherwise have included.

[6] Predictive quantile regressions can lead to quantile crossings (e.g., the 5% quantile being predicted above the 10% quantile). To address this issue, we estimate a fine grid of quantiles with $\theta \in \{0.05, 0.06, ..., 0.94, 0.95\}$ and apply the Chernozhukov et al. (2010) rearrangement method to en-

change in U.S. aggregate real personal consumption expenditures (excluding durables) over the past twelve months. The sample period covers January 1960 to September 2025. At each point in time, we pre-select the predictors and apply PQR to forecast out-of-sample at a 12-month horizon. The first out-of-sample prediction is for January 1970, and we extend the sample by one month at a time as new data become available.

2.2 Forecasting U.S. consumption growth: results

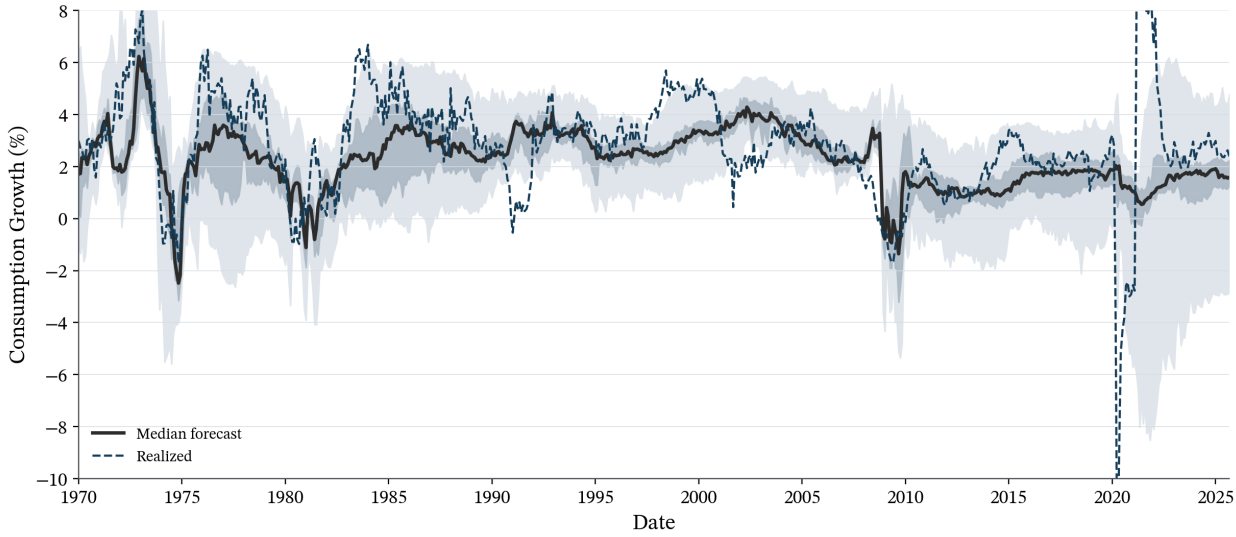
Figure 1 plots the predicted median (black) and the 5, 25, 75, and 95% quantiles (shades of grey) against the realized year-on-year growth rate of U.S. aggregate consumption (blue dashed). At each date, the blue dashed line is realized consumption growth over the *past* twelve months, while the quantiles are predictions for growth over the *next* twelve months using information available up to that date. The model captures the main business-cycle movements in consumption growth. Consumption volatility is high in the 1970s, and expected consumption growth falls during the recessions of the 1970s and 1980s. Realized consumption growth also falls in the 1990 and 2001 recessions, but forward-looking risks remain muted, consistent with how short and shallow these recessions were. Expected consumption growth begins a slow decline in the early 2000s but rises in the early stages of the 2008 recession even as realized consumption growth falls, possibly reflecting expectations of a quick recovery supported by Fed rate cuts and the 2008 tax rebates. The collapse of Lehman Brothers in September 2008 then produces an immediate fall in the mean and lower quantiles.

Since the Great Recession, expected consumption growth has been lower than before the crisis, and the distribution has become wider. The COVID-19 recession brought a dramatic fall in consumption growth, reflected in a decline of the 5% quantile to roughly −8.5% in late 2021. The median and upper quantiles fall only slightly. Tail risks were large, but the most likely outcome for growth over the next year was not expected to be as severe as in earlier recessions. After 2021, the lower quantile recovers as the shock fades, but the conditional distribution remains wider than before the pandemic, consistent with elevated downside risk through the 2022–2023 monetary-tightening cycle. By the end of the sample, the distribution has begun to narrow again as inflation moderates and recession fears recede.

Lower quantiles of the distribution are generally more volatile than upper quantiles, matching the evidence in Adrian et al. (2019) for the GDP growth distribution. Comove-

sure monotonicity of the quantile function. We evaluate forecasting performance only on the five targeted quantiles. To save computation time, we pre-select predictors via ℓ_1-penalized regression only for $\theta \in \{0.05, 0.15, 0.35, 0.5, 0.65, 0.85, 0.95\}$.

Figure 1: Predicted quantiles vs realized consumption growth, 12-month horizon



Note: The figure plots out-of-sample forecasts of the year-on-year U.S. real personal consumption expenditures (excluding durables) growth rate at a 12-month horizon. The black line is the estimated conditional median and the shaded grey regions are the estimated conditional quantiles, with the outer (lighter) band covering the 5–95% interval and the inner (darker) band covering the 25–75% interval. The blue dashed line is the realized growth rate of consumption over the *past* twelve months, while the predicted quantiles are for growth over the *next* twelve months using information available at each date. Quantiles are estimated by penalized quantile regression on factors extracted from the FRED-MD dataset and forecasts are generated recursively, extending the sample one month at a time. The evaluation window is 1/1970–9/2025.

ment across quantiles was stronger during the 1970s and early 1980s, after which the upper quantiles became more stable, so that recessions were associated with drops in the median and lower quantiles but stable upper quantiles. Apart from the episodes above, the outlook generally tracks recent consumption growth. The forecasting model also adjusts quickly as new data arrive: while the COVID-19 crisis was virtually impossible to predict a year in advance, the data arriving in March and April 2020 are quickly incorporated into the forecast and produce strong revisions to tail risks.

To evaluate forecasting performance formally, we compute the Koenker and Machado (1999) R^1, analogous to the standard R^2 measure used in OLS settings. For the θ-th quantile, $R^1(\theta)$ compares the sum of weighted prediction errors between the forecasting model and the historical sample average:

$$R^1(\theta) = 1 - \frac{\sum_t \rho_\theta\left(c_{t,t+12} - x_t'\hat{\beta}_{\theta,t}\right)}{\sum_t \rho_\theta\left(c_{t,t+12} - \mathbf{1}'\bar{\beta}_{\theta,t}\right)}, \tag{6}$$

Table 1: Forecast performance by quantile (R^1)

Model	Sample	$Q(0.05)$	$Q(0.25)$	$Q(0.50)$	$Q(0.75)$	$Q(0.95)$	Average
ML forecast	1970-2025	3.07	10.48	14.61	15.24	13.05	11.29
ML forecast	1981-2025	-3.95	1.96	8.73	12.23	8.87	5.57
NFCI forecast	1981-2025	-2.67	3.01	1.17	3.62	7.45	2.58

The table shows the out-of-sample forecast performance for the machine learning forecast and a simple forecast based on the NFCI for different quantiles and the average across quantiles. Forecasting performance is evaluated using R^1, which is reported in percent.

where $\rho_\theta(z) = z(\theta - \mathbb{I}(z < 0))$ is the quantile loss function, $\hat{\beta}_{\theta,t}$ are the coefficients estimated using the data available in period t, and $\mathbf{1}'\bar{\beta}_{\theta,t}$ is the historical sample average of the θ-th quantile. R^1 lies between 0 and 1 for in-sample regressions but can be negative out of sample if the forecasting model performs worse than the historical sample average. To evaluate goodness of fit globally, we average R^1 across the 5, 25, 50, 75, and 95% quantiles.

The machine learning forecasts achieve an OOS R^1 of 11.29% over the full 1970–2025 evaluation window. By comparison, the original Adrian et al. (2019) approach, which uses the Chicago Fed's NFCI as the sole predictor x_t in equation (2), yields an R^1 of 2.58% over the shared 1981–2025 sample. Evaluating the machine learning forecast over that same sample gives an R^1 of 5.57%, more than twice the NFCI benchmark. Table 1 reports these averages along with the R^1 for each separate quantile. Over the shared sample, the machine learning model outperforms the NFCI forecast at the median, 75%, and 95% quantiles, with the largest gain at the 75% quantile (12.23% vs. 3.62%) and a second sizeable gap at the median (8.73% vs. 1.17%). At the 5% and 25% quantiles, the NFCI forecast performs better (−2.67% vs. −3.95% at $Q(0.05)$ and 3.01% vs. 1.96% at $Q(0.25)$). A broad predictor set therefore contributes most around the middle and upper-middle of the conditional distribution, where consumption outcomes depend on macroeconomic information beyond financial conditions.

To check that this performance is not an artifact of one particular factor construction, Appendix A reports a twenty-model out-of-sample horserace using the same R^1 criterion as Table 1. The alternatives include principal-component models with one through eight factors, PQR factors built on the full predictor panel, PQR factors built on principal components, and the corresponding pre-screened variants, evaluated over the full 1970–2025 sample, a post-1970s subsample, and a lower-quantile exercise focused on $Q(0.05)$, $Q(0.25)$, and $Q(0.50)$.

PQR with ℓ_1 pre-selection delivers the highest average fit in all three exercises. Principal-component models are less stable: the first component is a competitive single

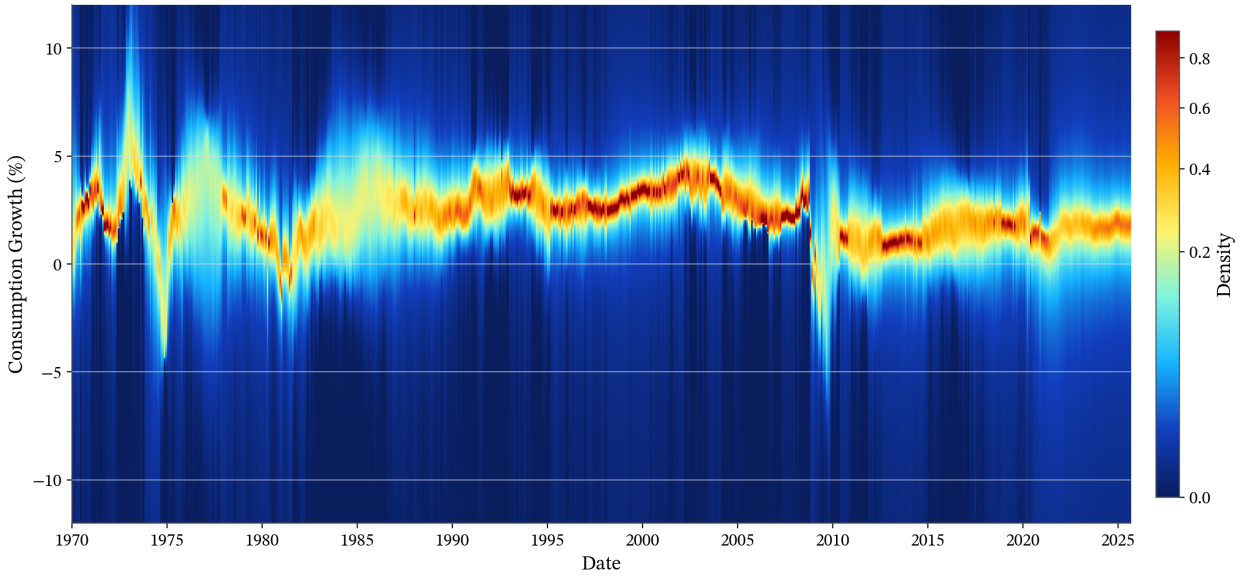
predictor, but adding untargeted components typically reduces and often reverses forecast accuracy. Pre-screening improves nearly every specification, indicating that the FRED-MD panel carries both informative variation and substantial redundancy. The preferred procedure therefore rests on three complementary ingredients: a broad macroeconomic information set, variable screening to remove weak predictors, and partial quantile regression to target the surviving factor at the quantile being forecast.

Taking the machine learning forecast as the benchmark, the NFCI-only forecast leaves a sizeable share of the variation in the middle and upper-middle of the conditional distribution unexplained, even though a richer predictor set can capture it. This may explain why NFCI-only forecasts in the literature tend to display even more stable interior quantiles than those reported here (Adrian et al., 2019; Loria et al., 2023). In the lower tail, by contrast, the NFCI is the more informative predictor, consistent with the finding of Adrian et al. (2019) that financial conditions carry information mainly about the lower tail of GDP growth. To the extent that an NFCI-only specification fits downside variation better than variation in the middle of the distribution, it overstates the asymmetry of the risk outlook, showing stronger fluctuations in downside relative to upside risk than are present in the data. Section 3 shows that this matters for return predictability: Kelley skewness from the NFCI-only forecast does not predict equity returns.

Appendix B complements these R^1 statistics by comparing realized year-ahead consumption growth with the corresponding predicted quantiles. The forecasts do not achieve exact coverage: empirical coverage is 11.3% for the 5% quantile and 14.2% for the 10% quantile, with larger shortfalls around the median and upper-middle quantiles. We therefore use the distribution as a time-varying description of relative tail shape rather than a set of exactly calibrated probability statements. That relative shape determines the Kelley skewness measure used below.

Appendix C reports which variables are selected most often for the consumption quantile forecasts. The selected variables span all groups of the McCracken and Ng (2016) dataset. Labor-market and interest-rate variables appear frequently across quantiles and over time. Prices matter for lower quantiles during the 1970s but much less afterward. Stock-market variables appear more often in upper-tail forecasts, although their role in lower-tail forecasts has increased over the sample. Housing variables are less prominent in upper-tail forecasts, whose selected predictors are also more stable over time. This variation across predictor groups helps explain why the model outperforms an NFCI-only forecast. Financial conditions contain useful information, but they do not exhaust the macroeconomic variation in future consumption risk. Consistent with this point, Bryzgalova et al. (2023) estimate that financial market innovations explain 25% of

Figure 2: Conditional density of year-ahead consumption growth over time



Note: The figure plots the fitted conditional density of year-on-year real personal consumption expenditures growth excluding durables over the next twelve months. For each month, the skewed t-distribution is fitted to the out-of-sample quantile forecasts and evaluated over possible future consumption growth outcomes on the vertical axis. Warmer colors indicate higher probability density, while darker blue regions indicate low density. The sample period is 1/1970–9/2025.

consumption variation.

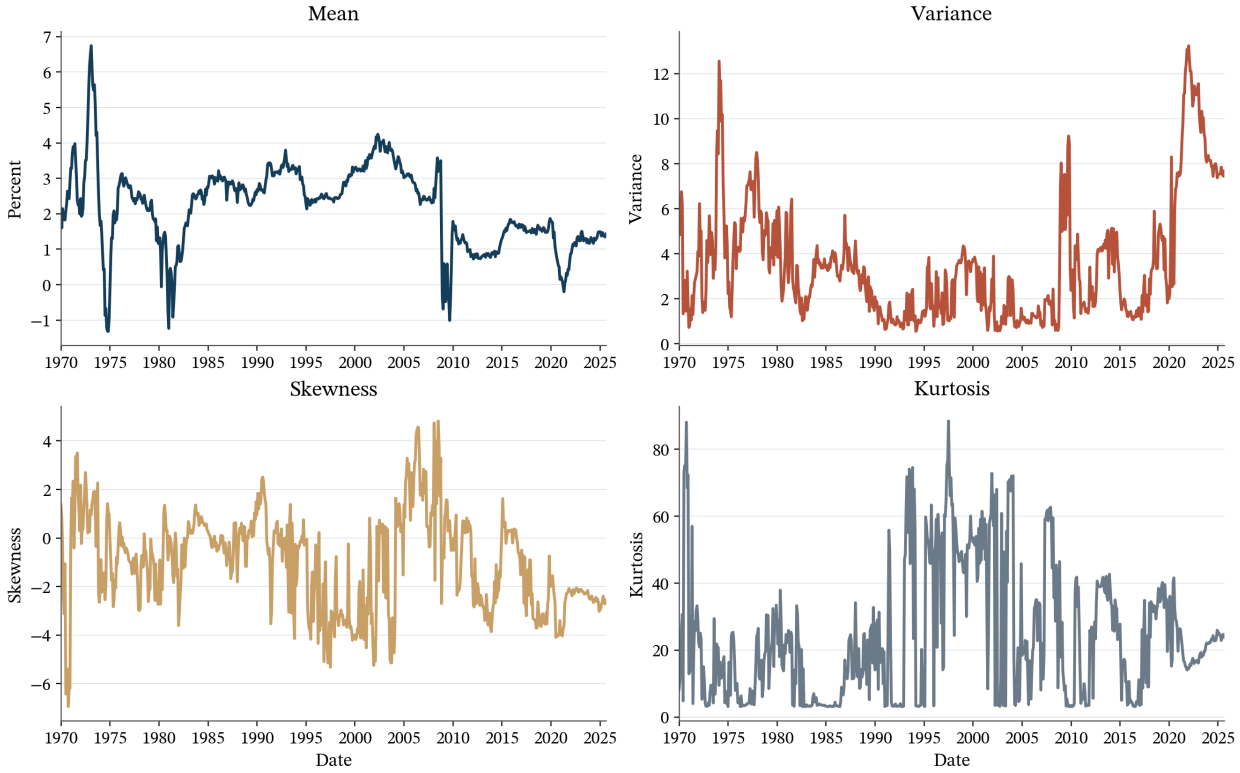
2.3 Approximating the consumption growth distribution

Given the estimated quantiles, we approximate the full conditional distribution of real consumption growth at each point in time. The fitted density is not needed to construct Kelley skewness, which is computed directly from the predicted 10th, 50th, and 90th percentiles below. We use the skewed t approximation to visualize the full conditional density and to construct additional distribution-based risk measures. Following Adrian et al. (2019), we fit the skewed t-distribution of Azzalini and Capitanio (2003) across the targeted quantiles in a given period:

$$f(c;\mu,\sigma,\alpha,\nu) = \frac{2}{\sigma} t\left(\frac{c-\mu}{\sigma};\nu\right) T\left(\alpha\frac{c-\mu}{\sigma}\sqrt{\frac{\nu+1}{\nu+\left(\frac{c-\mu}{\sigma}\right)^2}};\nu+1\right), \tag{7}$$

where $t(\cdot)$ and $T(\cdot)$ are the probability density function and cumulative distribution function of the Student-t distribution. We denote the PDF of the skewed t-distribution as $f(\cdot)$

Figure 3: Time-varying moments of the fitted year-ahead consumption growth distribution



Note: The figure reports the mean, variance, skewness, and kurtosis implied by fitting a skewed t-distribution to the out-of-sample quantile forecasts for year-on-year real personal consumption expenditures growth excluding durables over the next twelve months. The sample period is 1/1970–9/2025.

and its CDF as $F(\cdot)$.

The skewed t-distribution captures key properties of the consumption growth distribution without imposing too much structure on the data. It has four parameters: μ (location), σ (scale), α (shape), and ν (degrees of freedom, which determine tail thickness). As with the standard t-distribution, we can account for changes in location, scale, and tail size. Unlike the standard t-distribution, however, the skewed t-distribution can capture asymmetries by using the cumulative distribution function to shape the probability density function. It can therefore capture different behavior of upper and lower quantiles, reflecting differences between upside and downside risks in the conditional distribution. The shape parameter α governs the extent to which the CDF skews the PDF. For $\alpha = 0$, the distribution is symmetric. For $\nu \to \infty$, the skewed t-distribution converges to a skewed Gaussian distribution.

The skewed t-distribution in period t is fitted by choosing $\{\mu_t, \sigma_t, \alpha_t, \nu_t\}$ to minimize

the sum of squared differences between the estimated quantiles for $\{\theta_1, \theta_2, ..., \theta_J\}$ and the values implied by the skewed t-distribution:

$$\{\hat{\mu}_t, \hat{\sigma}_t, \hat{\alpha}_t, \hat{\nu}_t\} = \underset{\mu,\sigma,\alpha,\nu}{\operatorname{argmin}} \sum_{\theta} \left(\hat{Q}_{c_{t,t+h}|x_t}(\theta|x_t) - F^{-1}(\theta; \mu, \sigma, \alpha, \nu) \right)^2. \tag{8}$$

Figure 2 shows the fitted conditional density of year-ahead consumption growth over time. Each vertical slice is the probability density implied by the skewed t-distribution fitted to that month's predicted quantiles. The figure contains more information than the individual quantile paths in Figure 1: shifts in the center of the colored mass capture changes in the expected location of future growth, while changes in the width and asymmetry of the colored region capture time variation in uncertainty and tail risk. The density is relatively dispersed in the 1970s and early 1980s, tightens during much of the Great Moderation, and widens again around major downturns.

Recessions change the shape of the distribution as well as its center. During the Volcker disinflation and the Great Recession, the density moves lower and assigns more probability to weak consumption growth outcomes. The Great Recession is particularly persistent: after 2008, the density remains centered at lower growth rates and appears more diffuse than during the pre-crisis expansion, consistent with weaker growth prospects and greater macroeconomic uncertainty. The COVID-19 episode is different. The fitted density develops a pronounced lower tail, but the central and upper parts of the distribution move less than during the Great Recession. The fitted distribution therefore records the COVID-19 shock mainly as an increase in downside risk, not as a downward shift in all likely outcomes. These shape changes motivate measures such as Kelley skewness and left entropy, which distinguish left-tail risk from symmetric changes in volatility.

Figure 3 summarizes the same fitted densities by their first four moments. The mean and variance line up closely with the business-cycle interpretation of Figure 2. The fitted mean falls sharply during the early 1980s, the Great Recession, and the COVID-19 shock, and remains lower after the Great Recession than during much of the pre-crisis expansion. The variance rises when the predictive distribution widens, especially around the Great Recession and the COVID-19 episode.

The higher moments provide a less stable summary of the distribution. Skewness measures whether probability mass is tilted toward upside or downside outcomes, while kurtosis reflects the thickness of the tails. Both series move substantially over the sample, but their fluctuations are harder to map cleanly onto recession timing than movements in the mean and variance. This is consistent with Plagborg-Møller et al. (2020), who report similar findings for the GDP growth distribution. It also motivates our use of both

moments from the fitted distribution and quantile-based risk measures in the empirical analysis below. Quantile-based measures such as Kelley skewness and left kurtosis use specific points of the conditional distribution directly, preserving information about downside and upside risks without relying entirely on the fitted distribution's extreme-tail behavior.[7]

2.4 Constructing the risk measures

To remain agnostic about which measure of consumption risk best predicts returns, we construct several. The first group of measures is based on the predicted quantiles in Figure 1. We consider the 10%, 50%, and 90% quantiles individually and construct four additional risk proxies. *Dispersion* is the difference between the 90% and 10% quantiles, a standard measure in the uncertainty literature (e.g. Bloom et al., 2018). If uncertainty about future outcomes increases, dispersion should rise as the distance between high and low quantiles widens. *Kelley skewness* is defined as

$$ksk_t = \frac{\left(\hat{Q}_{c_{t,t+12}|x_t}(0.9) - \hat{Q}_{c_{t,t+12}|x_t}(0.5)\right) - \left(\hat{Q}_{c_{t,t+12}|x_t}(0.5) - \hat{Q}_{c_{t,t+12}|x_t}(0.1)\right)}{\hat{Q}_{c_{t,t+12}|x_t}(0.9) - \hat{Q}_{c_{t,t+12}|x_t}(0.1)}. \tag{9}$$

Kelley skewness is zero when the distance between the 90% quantile and the median equals the distance between the median and the 10% quantile. Negative values indicate a left-skewed distribution, since the 10% quantile is then further below the median than the 90% quantile is above it. Dividing by dispersion normalizes Kelley skewness to lie between -1 and 1. We use KSK as shorthand for Kelley skewness. Previous work has used Kelley skewness to document the procyclicality of skewness in the cross section of household (Busch et al., 2022) and firm (Salgado et al., 2023) outcomes. The last two quantile-based risk measures are *left kurtosis* and *right kurtosis* (Ferreira, 2018). Left kurtosis is defined analogously to Kelley skewness but uses the 45% and 5% quantiles centered around the 25% quantile, while right kurtosis compares the distances of the 95% and 55% quantiles from the 75% quantile.

The second group of risk measures is based on the fitted skewed t-distribution. We consider the mean, variance, skewness, and kurtosis shown in Figure 3. We also construct the two risk measures considered by Adrian et al. (2019). *Shortfall* measures downside

[7] We also do not construct separate distribution-based measures of consumption tail risk, which would be strongly affected by changes in the fourth moment.

risk as the expected value in the left part of the distribution up to the π quantile:

$$SF_t(\pi) = \frac{1}{\pi} \int_0^{\pi} \hat{F}^{-1}_{c_{t,t+12}|x_t}(\theta|x_t)d\theta. \tag{10}$$

During a recession, more probability mass shifts to below-median outcomes, lowering this left-tail mean; equivalently, the negative of shortfall would rise as a downside risk index. The analog for the right tail is *longrise*. *Left entropy* is:

$$\mathcal{L}^D_t(\pi) = -\int_{-\infty}^{\hat{F}^{-1}_{c_{t,t+12}|x_t}(\pi|x_t)} \left(\log \hat{g}_{c_{t,t+12}}(c) - \log \hat{f}_{c_{t,t+12}|x_t}(c|x_t)\right) \hat{f}_{c_{t,t+12}|x_t}(c|x_t)dc, \tag{11}$$

Left entropy measures the probability mass from a quantile with index π into the left tail of the conditional distribution f, relative to the probability mass that the unconditional distribution g assigns to those values. During a recession, stronger downside risks should be reflected in more probability mass in the left tail of the conditional distribution, raising left entropy. The right-tail analog, *right entropy*, integrates from π to infinity. We set $\pi = 0.5$ so that relative entropy compares the probability mass assigned to values above and below the median of the conditional distribution.[8] Relative entropy compares the conditional distribution to the unconditional distribution, whereas shortfall summarizes the conditional distribution in absolute terms. Skewness in the unconditional distribution therefore affects relative entropy but not shortfall.

To assess the business-cycle properties of the different risk measures, Table 2 reports their comovement with nine indicators of aggregate activity, uncertainty, and financial stress. The first two columns report correlations with past and future twelve-month consumption growth. The third column uses the future twelve-month excess return on the Fama-French market portfolio. The remaining columns use uncertainty or stress indicators: the Baker et al. (2016) economic policy uncertainty index, the Ludvigson et al. (2021) financial uncertainty index, the VIX, the variance risk premium, the TED spread, and the Moody's AAA corporate bond spread. These indicators all capture some aspect of aggregate risk but are not interchangeable. The VIX and TED spread are most directly tied to financial-market stress; LMN is a broader index of financial uncertainty; EPU reflects policy-related uncertainty and can move around events that are not purely macroeconomic; and the AAA spread captures credit-market conditions.

The table should be read with the sign of each consumption measure in mind. Some

[8]This makes the entropy measures distinct from tail-risk indices, since entropy does not focus on the extreme tail alone. Estimating extreme tails of macroeconomic variables suffers from large estimation uncertainty because macroeconomic disasters are rare. Entropy measured from the median, by contrast, is easier to estimate and does not rely on the extreme tails of the consumption growth distribution.

Table 2: Business-cycle correlations of consumption risk measures

	$c_{t-12,t}$	$c_{t,t+12}$	$r^M_{t,t+12}$	EPU_t	LMN_t	VIX_t	VRP_t	TED_t	AAA_t
Measures based on conditional distribution									
Mean	0.38	0.06	-0.29	-0.44	-0.17	0.01	0.00	0.22	-0.21
Variance	0.07	0.00	-0.03	0.35	0.36	0.12	0.03	-0.09	-0.06
Skew	0.04	-0.10	-0.17	-0.11	-0.18	-0.05	-0.10	0.21	-0.16
Kurtosis	-0.04	-0.03	-0.10	-0.23	0.03	0.06	0.08	0.04	0.13
Shortfall	0.25	0.03	-0.23	-0.47	-0.23	-0.03	-0.00	0.21	-0.10
Longrise	0.45	0.08	-0.32	-0.34	-0.05	0.08	0.01	0.26	-0.32
Left Entropy	-0.36	-0.10	0.08	0.21	0.22	0.11	0.04	-0.04	0.28
Right Entropy	-0.24	-0.09	0.06	0.07	0.14	0.13	0.16	-0.18	0.42
Measures based on quantiles									
$Q(0.1)$	0.13	-0.01	-0.17	-0.45	-0.20	0.00	0.01	0.20	-0.04
$Q(0.5)$	0.38	0.10	-0.25	-0.38	-0.17	-0.00	0.00	0.19	-0.23
$Q(0.9)$	0.45	0.04	-0.39	-0.39	-0.06	0.03	-0.00	0.30	-0.32
Dispersion	0.20	0.05	-0.11	0.33	0.20	0.02	-0.02	-0.04	-0.18
Kelley Skewness	0.07	-0.16	-0.31	-0.38	-0.06	0.03	-0.01	0.24	0.03
Left kurtosis	0.05	0.08	-0.02	-0.02	-0.17	0.00	-0.02	0.10	-0.16
Right kurtosis	-0.06	-0.07	-0.04	0.01	0.05	0.08	-0.05	0.20	-0.00
Measures based on Adrian et al. (2019) approach									
Mean	0.32	0.08	-0.17	-0.42	-0.26	-0.18	0.15	0.20	-0.34
Skew	-0.19	-0.19	-0.09	0.26	0.37	-0.00	-0.05	0.02	0.22
$Q(0.5)$	0.33	0.09	-0.17	-0.39	-0.29	-0.17	0.15	0.18	-0.37
Dispersion	-0.19	-0.02	0.00	0.41	0.32	0.55	-0.11	0.59	0.05
Kelley Skewness	-0.19	-0.11	0.08	0.00	0.18	-0.27	-0.02	-0.26	0.23

Each row in the table reports pairwise correlations between a risk measure and a business-cycle statistic. $c_{t-12,t}$ is year-on-year growth in real consumption excluding durables over the past 12 months. $c_{t,t+12}$ is consumption growth over the next 12 months. $r^M_{t,t+12}$ is the excess return on the Fama-French market portfolio between period t and $t+12$. EPU_t is the Baker et al. (2016) economic policy uncertainty index. LMN_t is the Ludvigson et al. (2021) financial uncertainty index at the 12-month horizon. VIX_t is the monthly average of a one-month VIX-type measure of S&P 500 implied volatility, computed from index option prices. VRP_t is the variance risk premium available on Hao Zhou's website. TED_t is the TED spread and AAA_t is the spread between AAA-rated corporate bonds and the yield on 10-year Treasuries as computed by Moody's. Correlations use the maximum available monthly overlap. For the machine learning rows, the first and third columns and LMN_t run from 01/1970 to 01/2025; $c_{t,t+12}$ runs from 01/1970 to 12/2024; EPU_t runs from 01/1985 to 01/2025; VIX_t runs from 01/1996 to 01/2025; VRP_t runs from 01/1990 to 12/2024; TED_t runs from 01/1986 to 01/2022; and AAA_t runs from 01/1983 to 01/2025. The bottom panel uses risk measures constructed from distribution forecasts using the Adrian et al. (2019) approach, which uses the NFCI as the sole predictor of future quantiles and is available from 01/1981 onward.

rows are growth-level summaries: a lower mean, lower quantile, lower shortfall, or lower longrise indicates worse expected consumption outcomes. Other rows are more naturally interpreted as risk or shape measures: higher variance, dispersion, and entropy indicate

more uncertainty or more probability mass away from the unconditional benchmark, while lower Kelley skewness indicates a more left-skewed conditional distribution. With this convention, the signs are economically sensible. The mean, median, upper quantile, shortfall, and longrise are positively correlated with past consumption growth and negatively correlated with EPU. For example, the fitted mean has correlations of 0.38 with past consumption growth and −0.44 with EPU, while the 90% quantile has correlations of 0.45 and −0.39 with the same two variables. These patterns indicate that the whole conditional distribution shifts down in weaker or more uncertain macroeconomic states.

The correlations with future consumption growth are small because the table reports unconditional monthly pairwise correlations, whereas the forecasting model conditions on a high-dimensional information set. Table 2 is a diagnostic of cyclical content; forecast performance is evaluated in Table 1. The strongest relationships are with contemporaneous state variables. Variance and dispersion are positively associated with EPU and LMN; left entropy is higher when past consumption growth is weak and financial uncertainty is high; and lower Kelley skewness coincides with higher EPU.

The bottom panel shows why a broad set of macroeconomic predictors matters. Consumption moments based on the Adrian et al. (2019) NFCI-only approach are more closely aligned with financial stress, especially for dispersion: the NFCI-only dispersion measure has correlations of 0.55 with the VIX and 0.59 with the TED spread. The corresponding machine learning dispersion measure is much less tied to these financial indicators, with correlations of 0.02 and −0.04, while still comoving with policy and financial uncertainty through EPU and LMN. This pattern is consistent with the evidence in Table 1: an NFCI-only forecast captures downside financial-stress states well but provides a narrower view of aggregate consumption risk than a model that draws on the wider macroeconomic panel.

In summary, the consumption risk measures show business-cycle correlations that are broadly in line with economic intuition. They are correlated with, but distinct from, existing measures of economic and financial risk. The correlations of Kelley skewness with the VIX and the variance risk premium are close to zero, and its correlation with LMN is only −0.06. Appendix A compares the forecasting performance of the machine learning and NFCI-only approaches in more detail. The next section tests whether the consumption risk measures predict equity returns.

3 Consumption risk and equity returns

This section studies whether consumption risk predicts stock market returns in the time series and in the cross section. The time-series tests compare Kelley skewness with the other consumption risk measures at several forecast horizons and under additional controls. Kelley skewness predicts excess market returns at all three horizons, while symmetric measures such as variance and dispersion do not. We then test whether Kelley skewness is priced across equity portfolios.

3.1 Consumption risk and the time series of returns

Let $r_{t,t+s}$ be the excess market return between period t and $t+s$. For a given forecasting horizon s, the predictive return regression is

$$r_{t,t+s} = a_s + b_s x_t + u_{t,t+s}, \tag{12}$$

where x_t is a scalar predictor. The dependent variable is the excess return on the Fama-French market portfolio, with the safe rate proxied by the one-month Treasury bill return. Both series are from Kenneth French's Data Library. The predictors are the consumption risk measures constructed in Section 2.4. All variables are in levels.

Table 3 reports univariate predictive regressions for forecasting horizons of 6, 12, and 36 months. Each row reports a separate regression with one consumption risk predictor, with standard errors corrected for overlapping returns using the Hodrick (1992) procedure. The strongest evidence comes from measures that capture the location or asymmetry of the conditional consumption growth distribution. The fitted mean, longrise, median, and 90% quantile are significant at the 6- and 12-month horizons, and the 90% quantile remains significant at the 36-month horizon. Their coefficients are negative: when the conditional distribution shifts lower or when its upper part deteriorates, future excess returns are higher. A one percentage point fall in the fitted mean, for example, is associated with a 4.05 percentage point increase in excess market returns over the next 12 months.

Kelley skewness is the most consistent predictor in Table 3. It is statistically significant at all three horizons, with coefficients of -6.93, -16.07, and -34.00 at horizons of 6, 12, and 36 months. A one-unit fall in Kelley skewness, such as from $+0.5$ to -0.5, is associated with a 16.07 percentage point increase in excess returns over the next 12 months. The sign aligns with a risk-premium interpretation: when the conditional distribution becomes more left-skewed and downside consumption risk rises, investors require higher expected returns

Table 3: Time series predictability - Univariate, in-sample

	Horizon h = 6			Horizon h = 12			Horizon h=36		
	Est.	SE	R^2	Est.	SE	R^2	Est.	SE	R^2
Measures based on conditional distribution									
Mean	-2.47*	0.96	6.3	-4.05*	1.75	8.5	-7.63	4.00	12.3
Variance	-0.12	0.42	0.1	-0.19	0.82	0.1	0.11	1.68	0.0
Skew	-0.98*	0.46	2.9	-1.35	0.85	2.8	-2.24	2.55	3.1
Kurtosis	-0.00	0.04	0.0	-0.09	0.09	1.1	-0.15	0.23	1.4
Shortfall	-1.48	0.80	3.4	-2.55	1.49	5.1	-5.67	3.23	10.1
Longrise	-2.92*	0.97	7.9	-4.70*	1.80	10.3	-8.31	4.51	13.4
Left Entropy	1.84	3.57	0.3	3.57	6.24	0.6	0.62	11.70	0.0
Right Entropy	7.93	4.36	2.1	4.46	7.47	0.3	4.53	18.50	0.1
Measures based on quantiles									
$Q(0.1)$	-0.74	0.60	1.5	-1.42	1.14	2.8	-3.77	2.48	7.7
$Q(0.5)$	-2.35*	0.99	5.4	-3.55*	1.78	6.2	-6.68	3.92	9.1
$Q(0.9)$	-2.90*	0.86	9.9	-5.05*	1.59	15.1	-8.26*	3.99	16.5
Dispersion	-0.77	0.70	1.1	-1.14	1.43	1.2	0.13	3.25	0.0
Kelley Skewness	-6.93*	2.91	3.5	-16.07*	5.45	9.5	-34.00*	16.67	16.3
Left kurtosis	-1.78	2.70	0.3	-0.79	4.97	0.0	-1.09	12.54	0.0
Right kurtosis	0.92	3.45	0.0	-2.87	5.63	0.2	-9.93	10.32	0.8
Measures based on Adrian et al. (2019) approach									
Mean	-3.14	3.32	1.1	-7.05	6.20	2.8	-14.80	14.10	5.2
Skew	-0.52	0.61	0.5	-0.96	1.19	0.9	-3.45	3.61	4.7
$Q(0.5)$	-2.60	2.52	1.4	-5.32	4.59	2.8	-7.97	10.65	2.6
Dispersion	-0.35	0.97	0.2	0.04	1.61	0.0	1.64	3.07	0.8
Kelley Skewness	4.15	5.85	0.4	7.36	11.30	0.6	-9.56	33.05	0.5
Measures of realized consumption									
Consumption	-0.97*	0.48	3.5	-2.25*	0.86	9.6	-3.22*	1.58	8.2
Consumption Lead	1.55*	0.48	9.2	2.19*	0.95	9.1	-0.86	1.68	0.6

The dependent variable is the excess return on the Fama-French market portfolio over the next $h \in \{6, 12, 36\}$ months. The predictor sample runs from 01/1970 until 07/2025 for $h = 6$, until 01/2025 for $h = 12$, and until 01/2023 for $h = 36$, with the last forecast made for returns over the h months after the final predictor observation. For the risk measures based on NFCI forecasts, the predictor sample starts in 01/1981. All data is at the monthly frequency. See the main text for details on the variable construction. All regressions include a constant, for which the coefficient estimate is not shown in the table to conserve space. The standard errors correct for overlapping returns as in Hodrick (1992). The asterisk (*) indicates statistical significance at the 5% level. The R^2 is reported in percent.

to hold equities. Its R^2 is among the largest at the 12- and 36-month horizons, exceeded at 36 months only by that of the 90% quantile. By contrast, the coefficient estimates on variance and dispersion, which summarize symmetric changes in uncertainty, are small and insignificant across horizons. While a large literature focuses on fluctuations in sym-

Table 4: Time series predictability - Univariate, in-sample, alternative samples

	1980–2025			1970–2006			1970–2025		
	Est.	SE	R^2	Est.	SE	R^2	Est.	SE	R^2
Measures based on conditional distribution									
Mean	-2.44	2.02	2.7	-3.58	2.51	5.9	-4.05*	1.75	8.5
Variance	0.21	0.89	0.1	-1.50	1.38	3.6	-0.19	0.82	0.1
Skew	-0.97	1.00	1.5	-0.37	0.89	0.2	-1.35	0.85	2.8
Kurtosis	-0.17	0.09	4.4	-0.02	0.09	0.1	-0.09	0.09	1.1
Shortfall	-1.69	1.64	2.3	-2.14	2.26	2.8	-2.55	1.49	5.1
Longrise	-2.42	2.34	2.0	-4.66	2.41	10.1	-4.70*	1.80	10.3
Left Entropy	-0.12	8.17	0.0	3.11	7.23	0.6	3.57	6.24	0.6
Right Entropy	-0.84	9.70	0.0	2.78	9.36	0.1	4.46	7.47	0.3
Measures based on quantiles									
$Q(0.1)$	-1.03	1.26	1.5	-1.05	1.65	1.2	-1.42	1.14	2.8
$Q(0.5)$	-1.87	2.09	1.5	-2.90	2.46	3.9	-3.55*	1.78	6.2
$Q(0.9)$	-3.72	2.15	5.5	-5.31*	2.25	14.9	-5.05*	1.59	15.1
Dispersion	-0.09	1.82	0.0	-1.85	1.68	3.0	-1.14	1.43	1.2
Kelley Skewness	-17.41*	6.92	10.1	-11.81*	5.69	4.8	-16.07*	5.45	9.5
Left kurtosis	4.09	5.49	0.7	2.24	5.80	0.2	-0.79	4.97	0.0
Right kurtosis	-6.88	7.25	0.8	0.42	6.17	0.0	-2.87	5.63	0.2
Measures based on Adrian et al. (2019) approach									
Mean	-7.05	6.20	2.8	-3.68	7.55	0.9	-7.05	6.20	2.8
Skew	-0.96	1.19	0.9	-1.62	1.28	3.5	-0.96	1.19	0.9
$Q(0.5)$	-5.32	4.59	2.8	-3.64	5.52	1.6	-5.32	4.59	2.8
Dispersion	0.04	1.61	0.0	-0.52	1.68	0.3	0.04	1.61	0.0
Kelley Skewness	7.36	11.30	0.6	6.09	12.07	0.7	7.36	11.30	0.6
Measures of realized consumption									
Consumption	-1.94*	0.92	7.5	-1.76	1.59	3.4	-2.25*	0.86	9.6
Consumption Lead	2.04	1.07	8.0	3.81*	1.52	16.1	2.19*	0.95	9.1

The dependent variable is the excess return on the Fama-French market portfolio over the next $h = 12$ months. Across the different columns, the requested predictor samples are 1/1980–1/2025 (left), 1/1970–12/2006 (middle), and 1/1970–1/2025 (right). For the risk measures based on NFCI forecasts, the actual sample starts in 1/1981 whenever the requested start date is earlier, so the left and right columns for the Adrian et al. (2019) approach use the same 1/1981–1/2025 sample. The last forecast is made for returns over the 12 months after the end of the predictor sample. All data is at the monthly frequency. See the main text for details on the variable construction. All regressions include a constant, for which the coefficient estimate is not shown in the table to conserve space. The standard errors correct for overlapping returns as in Hodrick (1992). The asterisk (∗) indicates statistical significance at the 5% level. The R^2 is reported in percent.

metric concepts of risk (Bloom, 2014), these estimates suggest that downside asymmetry is more informative for equity returns than symmetric uncertainty alone.

The final rows of Table 3 help separate conditional risk measures from realized con-

sumption growth. Past realized consumption growth predicts future excess returns with a negative coefficient: −2.25 and an R^2 of 9.6% at the 12-month horizon. Future realized consumption growth, by contrast, is positively associated with future excess returns at the 6- and 12-month horizons. The positive coefficient supports the use of conditional forecasts. If future realized consumption growth were part of investors' information set at time t, their required risk compensation should fall in response to higher growth expectations, so that excess returns should fall. To the extent that future risk compensation explains most equity return movements (Campbell and Ammer, 1993; Knox and Vissing-Jørgensen, 2022), it seems most likely that investors' forecasts of consumption growth differed from realized growth by enough to explain the positive coefficient.[9]

The bottom panel shows that not every conditional consumption forecast delivers the return predictability result. None of the measures constructed from the Adrian et al. (2019) NFCI-only quantile forecasts is statistically significant under Hodrick standard errors.[10] The NFCI-only mean and median have negative coefficients, but their R^2 values are modest, and NFCI-only Kelley skewness has the wrong sign at the 6- and 12-month horizons, with an R^2 of only 0.6% at 12 months.

Table 4 repeats the 12-month forecasting regressions for three sample windows to test robustness across subsamples. The first window starts in 1980 and removes most of the high-inflation 1970s; the second ends in December 2006 and excludes the Great Recession and the post-crisis period; the third is the full 1970–2025 baseline. Kelley skewness remains significant in all three windows. Excluding the 1970s leaves the coefficient at −17.41 with an R^2 of 10.1%. Ending the sample before the Great Recession weakens the fit, but the coefficient remains negative and significant at −11.81, with an R^2 of 4.8%, against 9.5% in the full sample. The result is therefore not a 1970s phenomenon, nor is it driven solely by the Great Recession.

Several other predictors are less stable across windows. The 90% quantile is significant in the pre-2007 and full-sample windows but not when the 1970s are excluded; the fitted mean, longrise, and median are significant only in the full sample. Symmetric risk measures remain weak: variance and dispersion are insignificant in all three windows. The NFCI-only rows again fail to reproduce the result: in both the pre-2007 and full-sample windows, NFCI-only Kelley skewness has the wrong sign at the 12-month horizon and very low fit.

Table 5 asks whether the Kelley skewness result survives controls for other predictors

[9] Alternative explanations would have to be based on lower-than-expected real rate realizations over the return horizon or stronger-than-expected dividend growth.

[10] Because the NFCI is only available from 1971 onwards, the sample for these risk measures starts in 1/1981 and runs to the same horizon-specific endpoints as the other predictors.

Table 5: Time series predictability - Bivariate, in-sample

	Skewness		Other Variable		
	Est.	SE	Est.	SE	R^2
Mean	-11.96*	5.78	-2.80	1.70	12.9
Shortfall	-13.50*	5.80	-1.16	1.48	10.3
Longrise	-12.55*	5.72	-3.77*	1.64	15.7
$Q(0.5)$	-14.05*	5.43	-2.79	1.61	13.1
$Q(0.9)$	-9.53	5.96	-4.11*	1.27	17.9
Consumption Lead	-13.97*	4.98	1.88*	0.72	16.1
LMN	-16.27*	5.64	-18.95	34.93	9.8
EPU	-12.20	7.85	8.16*	3.76	13.8
VIX	-16.86*	8.31	0.38	0.21	13.7
VRP	-15.67*	7.38	-0.01	0.05	10.0
TED	-15.10*	7.12	-3.19	4.42	9.6
AAA	-16.22*	6.95	0.51	3.70	9.8

The dependent variable is the excess return on the Fama-French market portfolio over the next $h = 12$ months. Each row reports a bivariate predictive regression using Kelley skewness and one additional predictor. The requested sample is 01/1970–01/2025. The actual row samples use the largest available common sample for each predictor pair: 01/1970–01/2025 for the consumption risk measures and LMN, 01/1970–12/2024 for Consumption Lead, 01/1985–01/2025 for EPU, 01/1996–01/2025 for VIX, 01/1990–12/2024 for VRP, 01/1986–01/2022 for TED, and 01/1983–01/2025 for AAA. The last forecast is made for returns over the 12 months after the end of the predictor sample. See the main text for details on the variable construction. All regressions include a constant, for which the coefficient estimate is not shown in the table to conserve space. The standard errors are Newey-West with 11 lags. The asterisk (*) indicates statistical significance at the 5% level. The R^2 is reported in percent.

one at a time. Each row reports a bivariate predictive regression for 12-month excess market returns using Kelley skewness and the additional variable named in the first column. The first group of controls contains the consumption-distribution measures most relevant in the univariate regressions, together with future realized consumption growth as a benchmark. The second group contains the six macro-financial risk indicators of Table 2. We use Newey-West standard errors with 11 lags rather than Hodrick (1992) standard errors, because the Hodrick correction is designed for the null of no predictability by any predictor and is not formally valid in these multi-predictor regressions (see footnote 38 in Bollerslev et al. (2015)).

Kelley skewness is negative in every specification and significant in ten of the twelve bivariate regressions. The estimates are also economically stable: outside the specifica-

tions that include the 90% quantile and EPU, the coefficient ranges from −11.96 to −16.86, close to the univariate 12-month estimate in Table 3. When the 90% quantile is included, the upper-tail quantile is itself significant and the Kelley skewness coefficient falls to −9.53, indicating some overlap between the upper-tail location of the forecast distribution and the asymmetric skewness measure. When EPU is included, EPU is significant and Kelley skewness loses significance, although its coefficient remains negative.

The remaining controls do not overturn the downside risk result. Longrise and future realized consumption growth are significant alongside Kelley skewness, raising the R^2 to 15.7% and 16.1%, respectively. LMN, VIX, VRP, TED, and AAA are not significant once Kelley skewness is included, while the Kelley skewness coefficient remains significant in each of those rows. The predictive power of Kelley skewness therefore does not come from the information in standard measures of uncertainty, volatility, the variance risk premium, or credit spreads.

The financial controls also relate the result to work that measures risk aversion, uncertainty, or tail-risk perceptions from asset prices (Bekaert et al., 2013; Hattori et al., 2016). Those measures capture the price of risk and market-implied tail risk, whereas Kelley skewness measures the quantity of downside risk in expected consumption.

A related concern is that consumption Kelley skewness simply captures existing growth-at-risk information. Appendix D benchmarks the measure against the New York Fed Outlook-at-Risk real-GDP series, a public implementation of the Adrian et al. (2019) growth-at-risk methodology. The two measures capture related but distinct variation, and in return predictability horse races the consumption measure retains its predictive content after controlling for the GDP-at-risk benchmark, while the benchmark itself does not.

Table 6 moves from in-sample regressions to a recursive forecasting exercise. At each forecast date, the model is estimated using only data available up to that date, beginning with the 01/1970–01/1980 training window. The estimation window then expands by one month, and the resulting forecast is compared with a recursively estimated unconditional benchmark that includes only a constant. A positive out-of-sample R^2 means that the predictor lowers mean squared forecast errors relative to the benchmark; a negative value means the benchmark performs better.

Kelley skewness has the strongest out-of-sample performance, with an out-of-sample R^2 of 7.78%. The 90% quantile also improves on the unconditional benchmark, but by a smaller margin of 3.10%. Future realized consumption growth produces only a modest gain of 1.35%, while longrise performs worse than the benchmark, with an out-of-sample R^2 of −5.65%. The ranking matches the in-sample evidence.

Table 6: Time series predictability - Out of sample

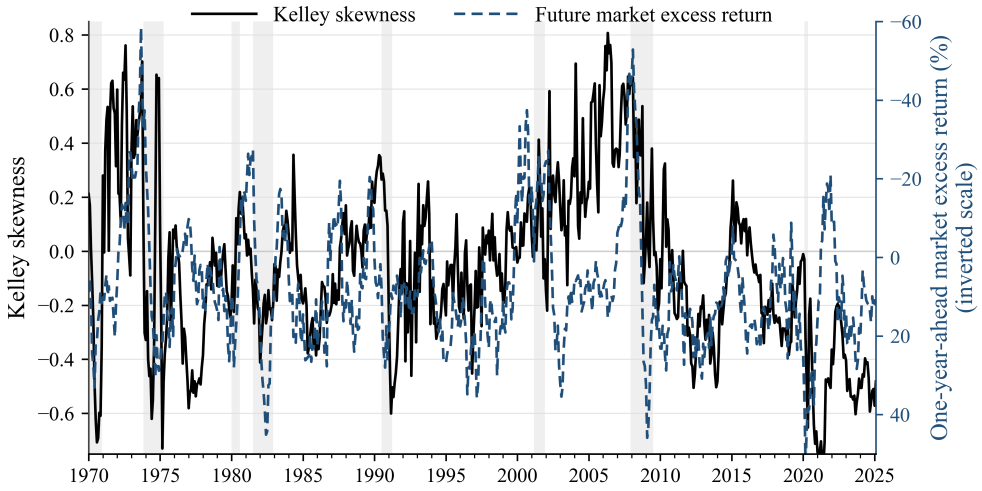
	OOS R^2	ENC-F stat.	Critical value
Kelley Skewness	7.78	96.49	147.40
$Q(0.9)$	3.10	97.11	179.02
Longrise	-5.65	35.23	107.81
Consumption Lead	1.35	121.80	261.25

The table reports results from expanding-window out-of-sample forecasting regressions. The dependent variable is the excess return on the Fama-French market portfolio over the next $h = 12$ months. The training sample begins in 01/1970. The first forecast uses predictor values in 01/1980 and is evaluated on the 12-month return ending in 01/1981. The final forecast uses predictor values in 01/2025 and is evaluated on the 12-month return ending in 01/2026. All regressions include a constant. The out-of-sample R^2 is reported in percent relative to a recursively estimated constant-only benchmark. The second column reports the Clark and McCracken (2005) ENC-F statistic for forecast encompassing of nested models, testing the null that the unconditional forecast has the same forecasting performance as the model in question. The 5% critical value is computed using a moving-block residual bootstrap with 1,000 repetitions and block length 12.

The statistical evidence is weaker than the economic magnitude. None of the ENC-F statistics exceeds its 5% bootstrap critical value, so the null of equal forecast performance against the constant-only benchmark is not rejected in the extended sample. We therefore interpret Table 6 as suggestive out-of-sample support rather than decisive forecast-encompassing evidence. Even so, an out-of-sample R^2 of 7.78% may seem small but is large in the context of return forecasting. Standard valuation ratios often deliver out-of-sample R^2 values below 2% or negative (Campbell and Thompson, 2008; Welch and Goyal, 2008), and the Kelly and Jiang (2014) tail-risk index reports an out-of-sample R^2 of 4.5% at the same horizon. A caveat is that the consumption risk estimates are produced recursively but use final revised macroeconomic data rather than real-time data vintages. To the extent that revisions improve the measured consumption risk forecasts, these results may overstate real-time forecasting performance.

Because Kelley skewness predicts returns across the in-sample specifications and has the strongest out-of-sample fit among the consumption-based predictors, we use it as the main consumption risk measure in the rest of the paper. Figure 4 plots Kelley skewness against future market excess returns. The right axis is inverted, so that high future returns appear low in the figure and move together with low Kelley skewness. Skewness falls during the 1973–1975 recession, the early 1980s, the 1990 recession, and the Great Recession, while future excess returns are elevated over the same periods.

Figure 4: Kelley skewness and future market excess returns



Note: The black line plots Kelley skewness, computed from the conditional distribution of real consumption growth excluding durables over the next 12 months. Lower values indicate greater left-skewness and hence higher downside consumption risk. The blue dashed line plots the cumulative excess return on the Fama-French market portfolio over the next 12 months. The right axis is inverted, so higher points on the blue line correspond to lower future excess returns. Grey shaded regions denote NBER recessions. The sample period is 01/1970–02/2025.

3.2 Consumption risk and the cross section of returns

In the cross section, assets with greater exposure to consumption risk should earn higher average excess returns. A standard prediction of a linearized consumption-based asset pricing model (Breeden et al., 1989; Kroencke, 2017) is that the year-on-year excess return is linear in the consumption betas $\beta_{c,i}$:

$$\mathbb{E}\left(r_{i,t,t+12}\right) = \lambda_{0,i} + \lambda_c \beta_{c,i}, \tag{13}$$

for asset i, where λ_c is the price of risk and $\lambda_{0,i}$ are asset-specific pricing errors that should be zero under the theory. We test this model using Fama and MacBeth (1973) regressions. Consumption betas for each asset i are estimated in a first stage as

$$r_{i,t,t+12} = \alpha_i + \beta_{c,i}\mathrm{ksk}_t + \epsilon_{i,t}, \tag{14}$$

where ksk_t is Kelley skewness as defined in (9). The second stage runs cross-sectional regressions in each period t:

$$r_{i,t,t+12} = \lambda_{0,i,t} + \lambda_{c,t}\beta_{c,i}. \tag{15}$$

Averaging over time gives the final estimates for the price of consumption risk and the pricing errors:

$$\lambda_c = \frac{1}{T}\sum_{t=0}^{T} \lambda_{c,t}, \quad \lambda_{0,i} = \frac{1}{T}\sum_{t=0}^{T} \lambda_{0,i,t}. \tag{16}$$

The theory in equation (13) predicts $\lambda_c < 0$ since low values of Kelley skewness indicate high consumption risk, and $\lambda_{0,i} = 0$.

Following Savov (2011), we omit the constant from the cross-sectional regressions, which helps identify the price of consumption risk when there is little variation in consumption betas across assets. Because the consumption betas are estimated, we apply the Shanken (1992) correction to the standard errors in the second stage. The test assets are the 25 Fama and French (1993) portfolios sorted on size and book-to-market ratio plus 12 Fama-French industry portfolios, giving 37 assets in total. All portfolio data are from Kenneth French's Data Library.

Table 7 shows that exposure to Kelley skewness is closely related to average returns across the 37 test assets. In the all-month specification, the cross-sectional price of risk is $\lambda_c = -0.51$ with a Shanken-corrected standard error of 0.08. The estimate has the sign predicted by the model: Kelley skewness is an inverse measure of downside consumption risk, so lower skewness corresponds to worse consumption risk states. The one-factor cross-sectional fit is high, with an R^2 of 90.6%, so exposure to the Kelley skewness factor explains a large share of average return differences across the test assets.[11]

The calendar-month rows check whether this result is driven mechanically by overlapping annual returns. The slope estimate is negative in every non-overlapping monthly subsample. It is statistically significant at the 5% level in April, June, July, August, September, and October, and remains close to significant in May. The estimates are less precise in January, February, March, November, and especially December, where the standard error is large and the cross-sectional R^2 falls to 48.2%. For the other calendar months, the R^2 remains above 80%. The estimated pricing errors (not reported) also show that the one-factor model does not price every asset exactly, which suggests that other risk factors are needed to span the full cross section. In summary, Kelley skewness of expected consumption growth is priced in the cross section and explains much, though not all, of the cross-sectional return variation.

[11] This cross-sectional test is conceptually distinct from reaching-for-income channels in which low rates raise demand for income-generating assets (Daniel et al., 2021). The test assets are sorted on size/book-to-market and industry, and the pricing variable is exposure to aggregate consumption skewness innovations rather than dividend yield or income demand.

Table 7: Fama-MacBeth cross-sectional pricing regressions

	λ_c	SE	R^2
All months	-0.51	0.08	90.6
January	-0.63	0.38	86.0
February	-0.49	0.28	85.8
March	-0.46	0.27	89.8
April	-0.47	0.23	91.4
May	-0.58	0.30	89.8
June	-0.49	0.23	90.7
July	-0.46	0.19	91.1
August	-0.38	0.15	91.1
September	-0.33	0.15	91.1
October	-0.45	0.21	90.0
November	-0.65	0.41	80.9
December	-0.97	0.93	48.2

The table reports results from Fama-MacBeth regressions over the 01/1970–02/2025 sample. The test assets are 25 Fama-French size/book-to-market portfolios and 12 Fama-French industry portfolios. The first row uses all year-on-year return observations and therefore contains overlapping return windows. The remaining rows use non-overlapping annual returns by keeping only one calendar month at a time, such as January-to-January returns. Standard errors for the slope estimate λ_c use the Shanken (1992) correction for generated first-stage betas. The R^2 is the cross-sectional fit and is reported in percent.

4 Shock drivers of consumption downside risk

Which shocks move consumption downside risk? The asset pricing results show that Kelley skewness is priced, but not what drives it. Because equity premia are high when Kelley skewness is low, the shocks that lower it point to the macroeconomic risks investors demand compensation for. They also point to the policies that might stabilize those risks.

We consider four non-policy candidates from the macro-finance literature. Tighter financial conditions and higher uncertainty have been linked to greater macroeconomic

tail risk (Adrian et al., 2019; Ludvigson et al., 2021; Adrian et al., 2022). Adverse oil-supply news (Känzig, 2021) and adverse news about investment-specific technology (Ben Zeev and Khan, 2015) can lower expected productive capacity. We estimate local projections for each shock and rank the shocks by the average response of Kelley skewness. We then turn to monetary policy, whose association with Kelley skewness depends on how high downside risk is when the surprise occurs.

4.1 Empirical design

For each shock j, we estimate quarterly local projections à la Jordà (2005):

$$\text{ksk}_{t+h} = \alpha_h^j + \beta_h^j s_t^j + \sum_{\ell=1}^{2} \Gamma_{h,\ell}^{j\prime} X_{t-\ell} + u_{t+h}^j, \tag{17}$$

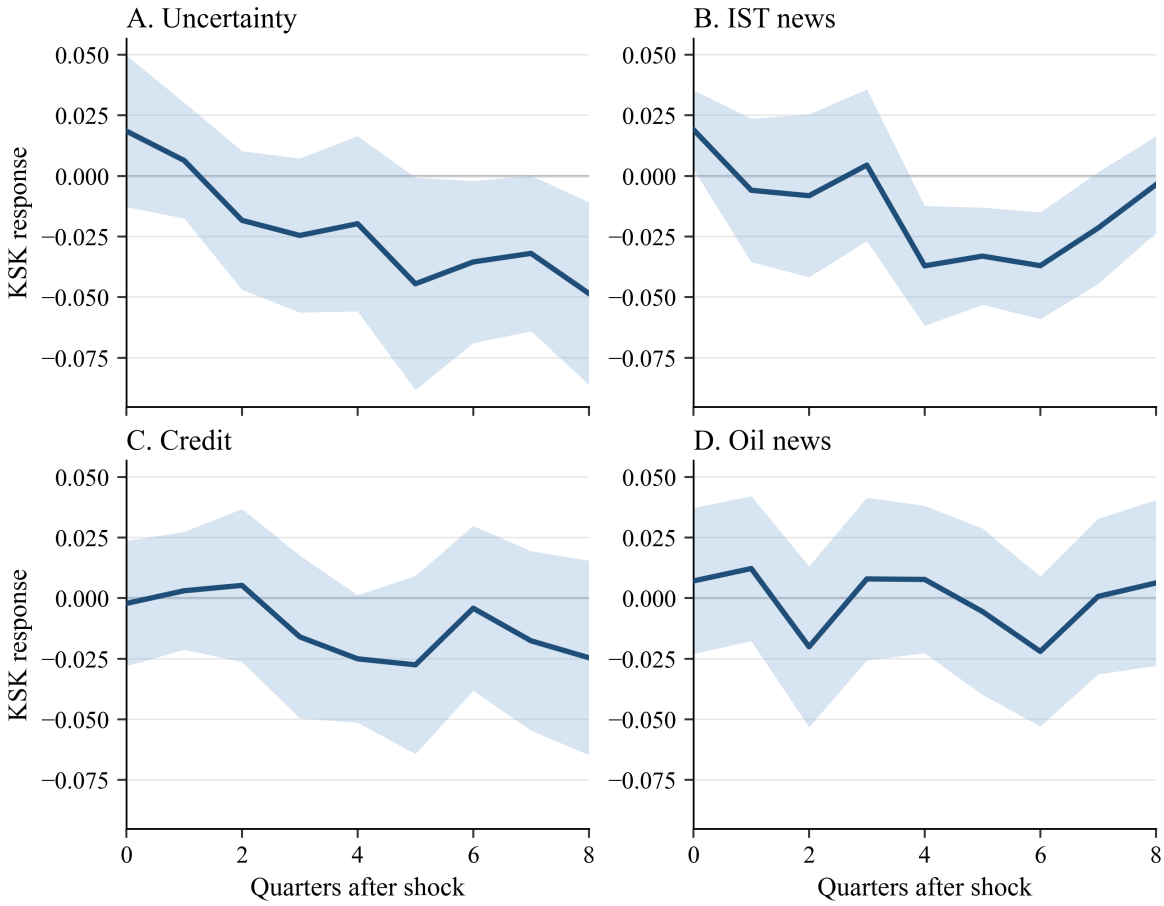
where ksk_{t+h} is Kelley skewness in the conditional distribution of year-ahead consumption growth and s_t^j is a standardized shock or innovation. We use quarterly local projections so that all four shocks are compared at the quarterly frequency of the investment-specific technology (IST) news series. For these quarterly LPs, monthly Kelley skewness and predicted quantiles are taken from the final month of each quarter. We adopt an adverse sign convention throughout: positive values denote tighter credit, higher financial uncertainty, adverse oil-supply news, and adverse IST news. A negative β_h^j therefore means that an adverse realization of shock j is associated with a more left-skewed distribution of expected consumption growth.

The control vector X_t contains Kelley skewness, industrial-production growth, CPI inflation, the Wu and Xia (2016) shadow rate, and the excess bond premium, each entering with two quarterly lags. Each shock series is standardized within the common estimation sample, so all reported responses correspond to a one-standard-deviation adverse realization in that window. Inference uses Newey-West standard errors with horizon-specific lag length.[12]

The shock set combines two externally identified news measures with two index innovations. Oil-supply news is the Känzig (2021) shock, aggregated from monthly to quarterly frequency by summing monthly innovations within each quarter, so that the quarterly shock captures the cumulative news arriving during the quarter before standardization. Adverse IST news uses the extended Ben Zeev and Khan (2015) series,

[12] Treating the estimated skewness index as observed understates full sampling uncertainty, so the reported confidence bands are likely too narrow. Re-estimating the out-of-sample quantile forecasts inside the LP bootstrap is computationally infeasible because the variable-selection and distribution-fitting steps take several hours.

Figure 5: Kelley skewness responses to broad aggregate shocks



Note: The figure reports local-projection responses of Kelley skewness to one-standard-deviation adverse shocks. Positive shock realizations correspond to tighter credit conditions, higher financial uncertainty, adverse oil-supply news, and adverse investment-specific technology news. The sample period is 1991Q1–2025Q2 for current shock values. Shaded bands are 90% Newey-West confidence intervals.

sign-flipped so that positive values are adverse. Credit conditions are AR(12) innovations in the excess bond premium of Gilchrist and Zakrajšek (2012) and Favara et al. (2016), and financial uncertainty is the analogous AR(12) innovation in the Ludvigson et al. (2021) financial uncertainty index.

4.2 Which shocks matter?

Figure 5 plots the impulse responses of Kelley skewness to each shock in the sample for which all four broad shocks are simultaneously available, 1991Q1–2025Q2 for current shock values.

Table 8: Broad-shock ranking for Kelley skewness

Shock	Avg. beta h2-h6	Avg. partial R2
Financial uncertainty	-0.029	0.014
Adverse IST news	-0.022	0.016
Credit conditions	-0.014	0.004
Oil supply news	-0.006	0.007

Note: The table ranks shocks over the 1991Q1–2025Q2 sample, for current shock values. Avg. beta h2–h6 is the average local-projection coefficient over horizons two through six quarters. Avg. partial R^2 averages the shock's partial R^2 over horizons two, four, and six. Positive shocks are adverse, so negative responses indicate a deterioration in Kelley skewness.

All four shocks lower Kelley skewness at medium horizons. Most paths are flat or close to zero in the first quarter, deepen between horizons two and six, and then revert toward zero. Financial-uncertainty innovations produce the deepest and most persistent decline, with confidence bands separating from zero over a sustained range of horizons. Adverse IST news and tighter credit conditions produce shallower declines, while oil-supply news barely moves Kelley skewness. Averaging over horizons two through six quarters gives the same ordering: financial uncertainty (-0.029), adverse IST news (-0.022), tighter credit conditions (-0.014), and oil-supply news (-0.006). These responses equal about 9, 7, 4, and 2 percent of the standard deviation of Kelley skewness in the common quarterly sample.

Table 8 pairs the same average response with the average partial R^2 at horizons two, four, and six. The partial R^2 is a descriptive incremental-fit statistic: it measures the share of variation in Kelley skewness that each shock explains beyond what the macroeconomic controls already absorb, rather than a structural decomposition of historical movements in downside risk. The partial-R^2 ranking is similar to the average-response ranking but not identical: adverse IST news has the largest value, followed by financial uncertainty, oil-supply news, and credit conditions. We rank shocks by their average response, which measures how far each shock moves Kelley skewness, and treat the partial R^2 as a descriptive complement.

Financial uncertainty produces the largest response. Higher uncertainty raises the value of liquidity, can delay investment and durable spending through real-options effects (Bernanke, 1983; Bloom, 2009; Bachmann and Bayer, 2014; Bloom et al., 2018), and increases precautionary saving and pressure on risk-bearing capacity (Jurado et al., 2015; Basu and Bundick, 2017; Ludvigson et al., 2021). These mechanisms can extend the lower tail of expected consumption growth without moving the center one-for-one. The result

also accords with Hyland (2026), who links financial-uncertainty shocks to the tail risks confronting monetary policy.

Adverse IST news produces the second-largest response. Such news signals lower future productive capacity (Beaudry and Portier, 2006; Fernald, 2014; Ben Zeev and Khan, 2015); as expected capacity contracts, the conditional growth distribution tilts downward and assigns more weight to bad states. Credit-condition innovations operate through a different set of mechanisms, including the financial accelerator (Bernanke et al., 1999), credit-cycle dynamics (Kiyotaki and Moore, 1997), and intermediary risk-bearing constraints (Adrian and Shin, 2010; Brunnermeier and Sannikov, 2014; Gertler and Karadi, 2015). Across the four shocks, the largest Kelley skewness responses are tied to uncertainty, expected capacity, and financial frictions. Models with balance-sheet constraints, investment irreversibility, or precautionary motives generate this type of asymmetric propagation (Brunnermeier and Sannikov, 2014; Basu and Bundick, 2017; Baqaee and Farhi, 2019). The response to credit conditions also agrees with evidence on the role of credit costs and risk-bearing constraints in macro-financial propagation (Gertler and Karadi, 2015; Favara et al., 2016; Drechsler et al., 2018). Here, however, the outcome is the induced asymmetry of expected consumption growth rather than the credit spread itself.

4.3 The quantile mechanism

Kelley skewness is a nonlinear function of the 10th, 50th, and 90th percentiles of the conditional distribution, which we denote q_{10}, q_{50}, and q_{90}. A decline can reflect a lower tail that falls further than the median, an upper tail that compresses toward the median, or both. If the priced object reflects asymmetric movements in the lower tail, the shocks most closely associated with it should reduce q_{10} more than q_{50}.

We re-estimate equation (17) using the predicted quantiles and the lower-tail quantile difference as outcomes. Figure 6 plots the responses for the three strongest shocks in the broad-shock ranking. Oil-supply news is omitted from the figure because it produces the smallest Kelley skewness response in Table 8. Table 9 reports the average responses over horizons two through six quarters.

Averaged over horizons two through six quarters, financial-uncertainty innovations move the 10th percentile of expected consumption growth by −0.283, against −0.156 at the median, so the gap between the median and the 10th percentile widens by 0.127. Uncertainty therefore worsens bad outcomes more than central ones instead of acting as a pure level shock, consistent with real-options and intermediary risk-bearing channels (Bloom, 2009; Brunnermeier and Sannikov, 2014). Adverse IST news shows the same pattern with smaller magnitudes: q_{10} falls by 0.128 against 0.043 at the median, so the

Table 9: Quantile mechanism for the leading shocks

Shock	KSK	q10	q50	q10-q50
Adverse IST news	-0.022	-0.128	-0.043	-0.085
Credit conditions	-0.014	-0.047	-0.067	0.019
Financial uncertainty	-0.029	-0.283	-0.156	-0.127

Note: The table reports average responses over horizons two through six quarters. The $q_{10} - q_{50}$ column compares the 10th-percentile response with the median response; more negative values indicate that the lower tail falls by more than the center of the distribution.

gap widens by 0.085. News of lower future productive capacity thus lowers the median modestly but the lower tail substantially.

For credit-condition innovations, the quantile pattern is different. The 10th-percentile response is -0.047 and the median response is -0.067, so the gap between the median and the 10th percentile narrows by 0.019. Credit tightening therefore looks like a broad downward shift in the conditional distribution: bad and central outcomes both deteriorate, but the relative gap closes slightly. A credit channel operating through aggregate demand and balance-sheet repair would depress the whole distribution in this way. This pattern also explains why credit conditions move Kelley skewness less than financial uncertainty or IST news despite their central role in macro-financial propagation.

In summary, financial uncertainty and adverse IST news stretch the lower tail of expected consumption growth, which is the feature that Section 3 finds to be priced in equity markets.

4.4 A closer look at monetary policy

Monetary policy is a stabilization instrument, so its effect on downside risk may depend on how high that risk already is. Easing may reduce downside risk most when it is elevated, because the lower tail of expected consumption growth then has the most room to recover. When downside risk is low, risk-taking and vulnerability-building channels can mute or even reverse the later effects of accommodation (Borio and Zhu, 2012; Schularick and Taylor, 2012; Brunnermeier and Sannikov, 2014). An average response may therefore combine opposing effects across states.

Because high-frequency monetary surprises are available monthly, we estimate monthly local projections of Kelley skewness on monetary surprises, which give more observations than the quarterly panel above. We allow the response to depend on $\widehat{\text{ksk}}_t$, the six-month moving average of recent Kelley skewness, and evaluate it at $\widehat{\text{ksk}}_t = -0.4$

and +0.4. We refer to these as the low-KSK state, in which downside risk is high, and the high-KSK state. The benchmark surprise is the Jarociński and Karadi (2020) monetary-policy component, which separates pure policy news from the central-bank information component. Appendix E reports robustness to alternative surprise series, sample windows, and state-indicator definitions.

The unconditional response is small. Figure 7 plots the average Kelley skewness response to a one-standard-deviation monetary easing; averaging over horizons six through eighteen months yields 0.0016 index points, under five percent of the in-sample standard deviation of Kelley skewness. The state-conditioned responses below are much larger and have opposite signs.

The weak average estimate hides substantial heterogeneity across states. Figure 8 plots the baseline and interaction coefficients from the state-dependent specification together with the two implied response paths. The interaction term is negative at medium horizons, so the estimated response to easing is larger when recent Kelley skewness is low. Averaged over horizons six through eighteen months, the response is +0.0496 in the low-KSK state and −0.0224 in the high-KSK state, a wedge of 0.0721 index points, against an unconditional response of 0.0016. The small unconditional response therefore averages two opposing effects.

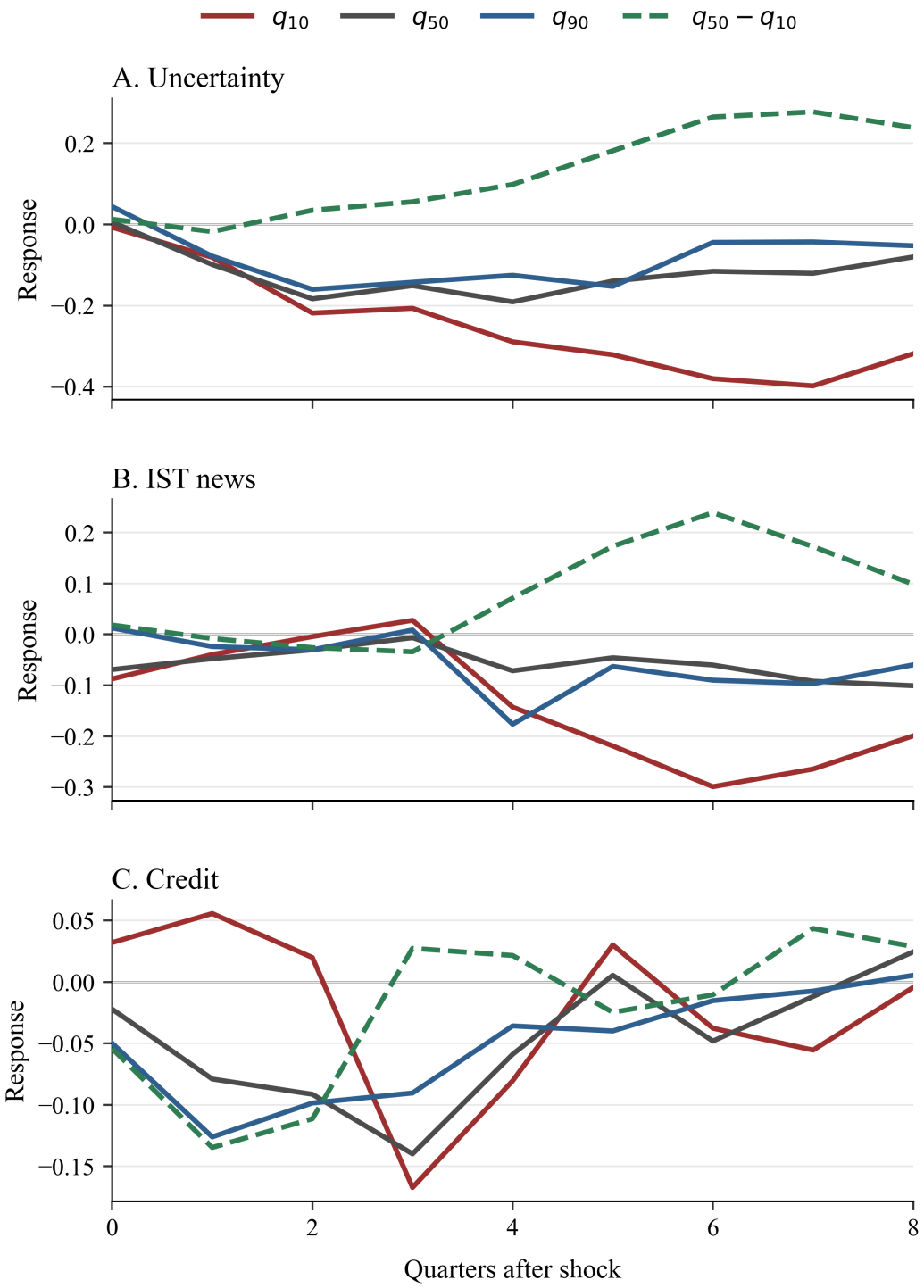
The result concerns a narrower channel than the broader literature on monetary policy and equity prices. Event-study and high-frequency studies estimate the total asset-price response to policy surprises, including cash-flow news, discount-rate news, credit-cost effects, term-premium movements, and investor demand for income-generating assets (Bernanke and Kuttner, 2005; Gertler and Karadi, 2015; Hanson and Stein, 2015; Daniel et al., 2021). We instead estimate how policy surprises are associated with the conditional left tail of expected consumption growth. The muted average Kelley skewness response means that this channel is concentrated in the low-KSK state; it does not imply that monetary policy has small effects on asset prices overall.

Quantile responses show that this state dependence operates mainly through the lower tail. Figure 9 plots the responses of q_{10}, q_{50}, and q_{90} in the low-KSK state, together with the directly estimated response of the left gap $q_{50} - q_{10}$. Over horizons six through eighteen months, easing raises the 10th percentile by 0.1296, against 0.0384 at the median and 0.0952 at the 90th percentile. The lower tail therefore rises by 0.0913 more than the median, and the directly estimated left gap narrows by a similar 0.0749. In other words, following an easing in the low-KSK state, the distance between the median and poor outcomes falls relative to the distance between good and median outcomes. Appendix E shows that this lower-tail mechanism is not unique to the benchmark Jarociński and Karadi (2020)

surprise: several alternative high-frequency surprise series also pair positive low-KSK Kelley skewness responses with narrower left gaps.

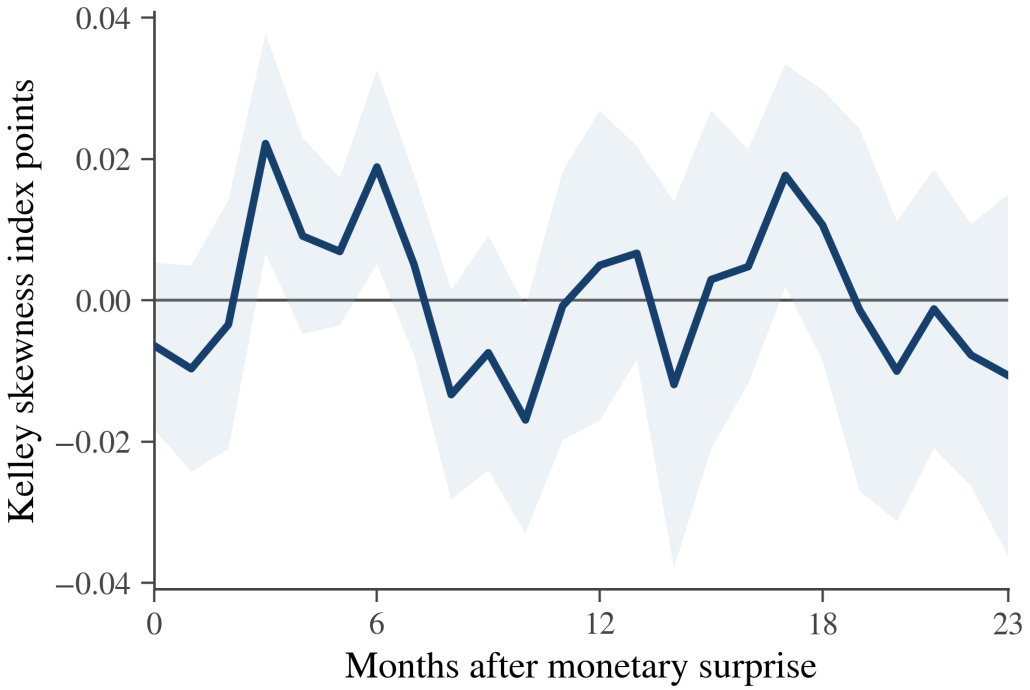
Earlier work shows that monetary policy moves option-implied tail risk and financial risk aversion (Bekaert et al., 2013; Hattori et al., 2016). Our evidence concerns the quantity of risk rather than its price: in the low-KSK state, policy surprises lift the lower quantiles of expected consumption growth. In summary, these findings are consistent with both the idea that loose financial conditions in good times can build vulnerabilities (Schularick and Taylor, 2012; Brunnermeier and Sannikov, 2014) and the idea that easing can mitigate downside risks to near-term growth in crisis times. The association is concentrated in the states in which expected equity returns are highest.

Figure 6: Quantile responses for the leading shocks



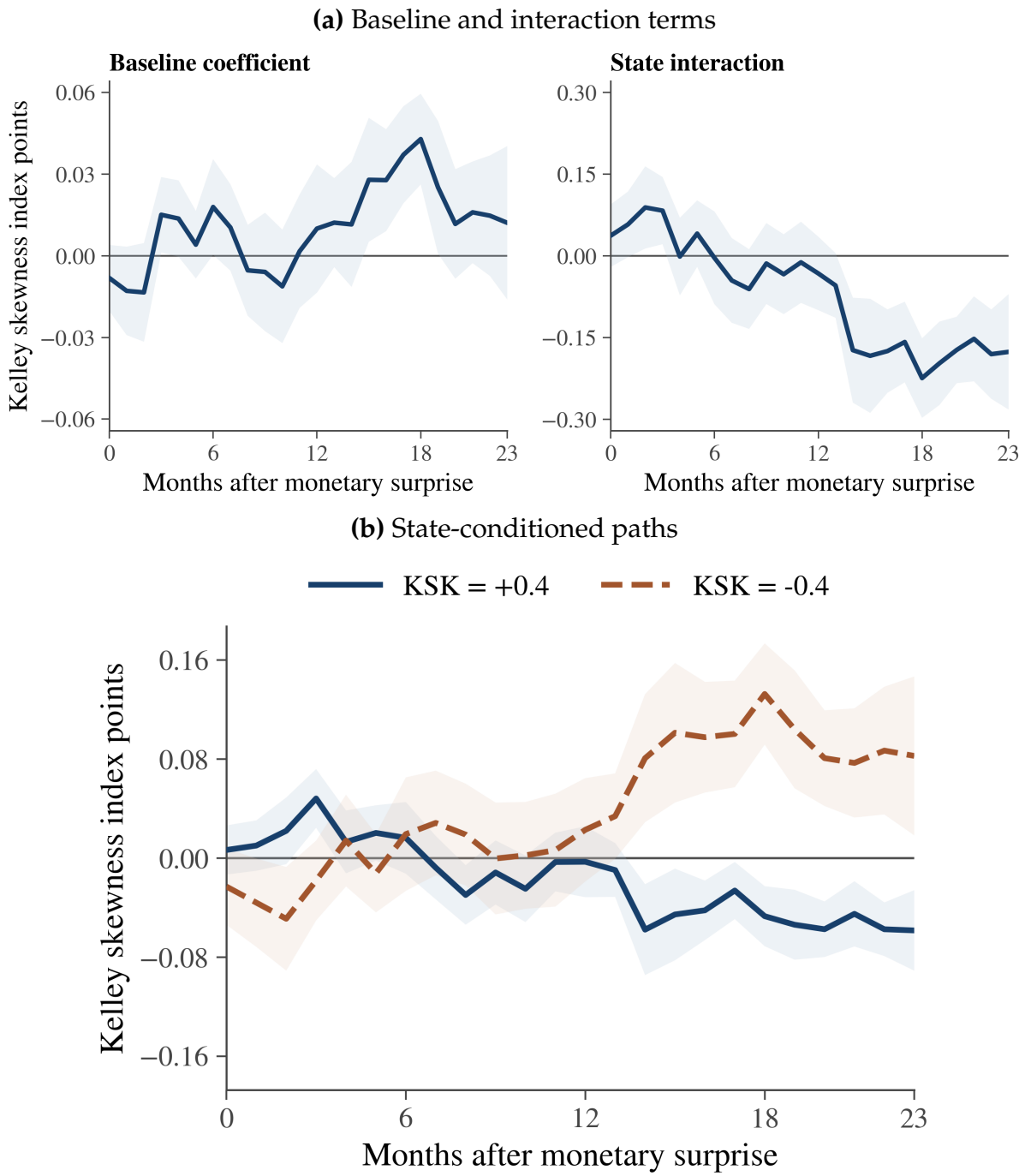
Note: The figure reports local-projection responses of the 10th, 50th, and 90th predicted quantiles of year-ahead consumption growth, together with the difference $q_{50} - q_{10}$. Positive shocks are adverse. The sample period is 1991Q1–2025Q2 for current shock values.

Figure 7: Average Kelley skewness response to monetary easing



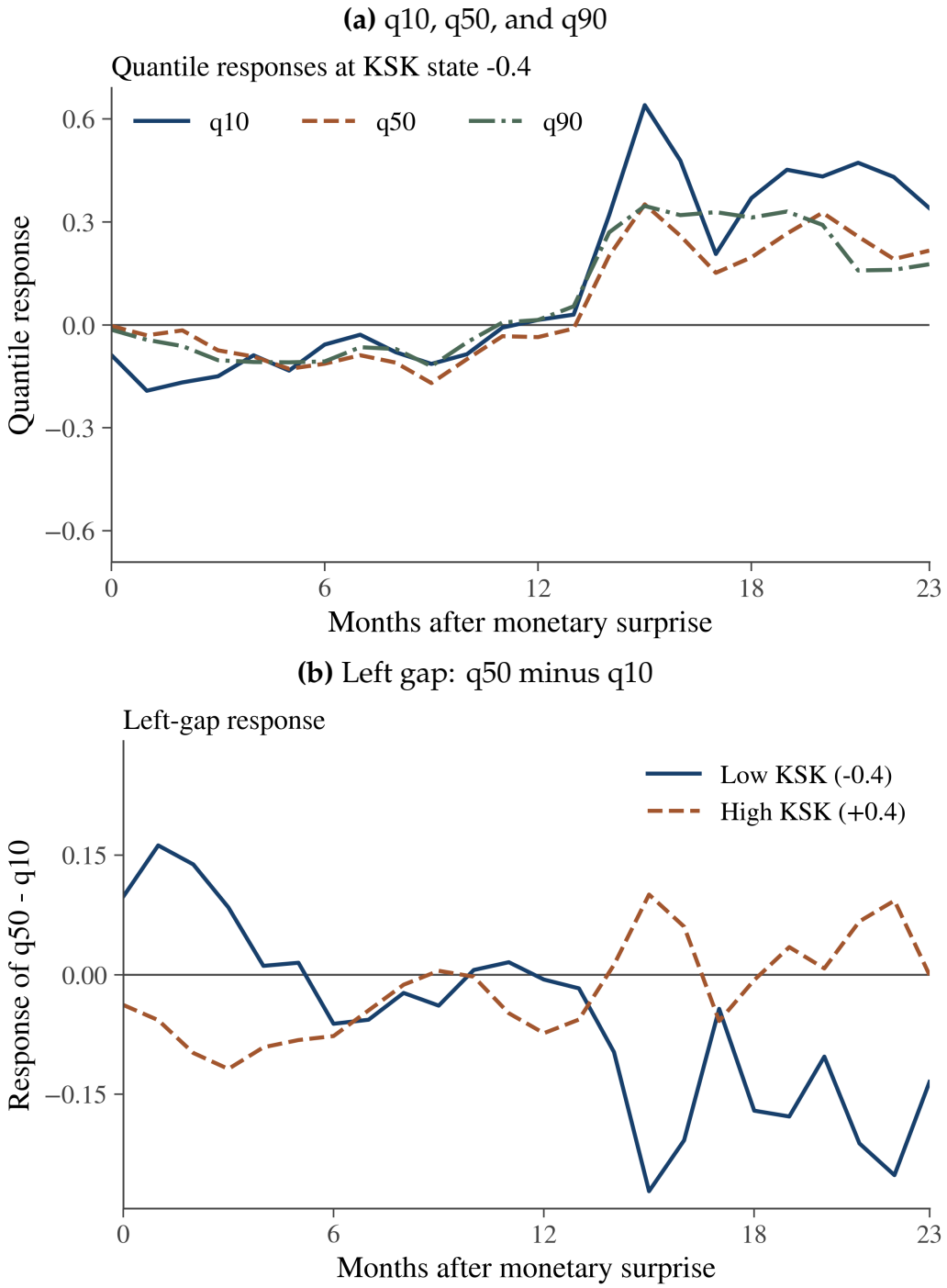
Note: The figure plots the response of Kelley skewness to a one-standard-deviation monetary easing using the benchmark Jarociński and Karadi (2020) monetary-policy component. The sample period is 1994-02–2025-06. The controls include log industrial production, log CPI, the shadow rate, and the excess bond premium, as well as lags of Kelley skewness and monetary surprises. All data are monthly and the lag length is 12 months. Shaded bands are 90% Newey-West confidence intervals.

Figure 8: State dependence in the Kelley skewness response to monetary easing



Note: Panel (a) plots the baseline response coefficient and the state interaction term using the benchmark Jarociński and Karadi (2020) monetary-policy component. Panel (b) evaluates the implied response at recent Kelley skewness equal to +0.4 and −0.4, where recent skewness is measured as a six-month moving average. Panel (b) shows that the small unconditional response masks opposite state-contingent responses: easing raises Kelley skewness in low-KSK (high-downside-risk) states but has a weaker or opposite effect in high-KSK states. The sample period is 1994-02–2025-06. All data are monthly and the lag length is 12 months. Shaded bands are 90% Newey-West confidence intervals.

Figure 9: Quantile responses to monetary easing in low-KSK states



Note: Panel (a) shows state-conditioned responses of the 10th, 50th, and 90th predicted quantiles of year-ahead consumption growth when recent Kelley skewness equals −0.4, the low-KSK (high-downside-risk) state. Panel (b) shows the response of $q_{50} - q_{10}$ for low- and high-KSK states. A negative response of $q_{50} - q_{10}$ indicates that the lower tail improves relative to the median. The monetary surprise is the benchmark Jarociński and Karadi (2020) monetary-policy component. The sample period is 1994-02–2025-06. All data are monthly and the lag length is 12 months.

5 Conclusion

This paper estimates the conditional distribution of future consumption growth, constructs risk measures based on it, and uses these to study the link between expected consumption risk and excess equity returns. We find that Kelley skewness, a measure of downside relative to upside risk in consumption growth, is priced in the time series and the cross section of returns. Stronger downside risks predict higher future excess returns, in line with the standard theory. In contrast, we find no comparable predictive power for the variance or dispersion of the distribution. Our analysis also points to the importance of drawing on a broad set of macroeconomic variables when estimating consumption risk. Existing work such as Adrian et al. (2019) uses financial conditions as the only predictor, and Kelley skewness estimated in this way does not predict returns.

To understand the drivers of consumption risk, we then study how it responds to macroeconomic shocks. Financial uncertainty and adverse investment-specific technology news lower Kelley skewness the most, mainly by stretching the lower tail of expected consumption growth, while credit tightening shifts the whole distribution down. Monetary easing is associated with lower downside risk in high-risk states and with higher downside risk in low-risk states. These state-dependent associations suggest that loose monetary policy may help manage crises while also contributing to the buildup of vulnerabilities before a downturn.

For asset pricing theory, the results imply that consumption-based models need to generate time variation in the asymmetry of expected consumption growth, not only in its level and variance. New Keynesian vulnerability models show how monetary policy can affect the conditional distribution of macroeconomic outcomes (Adrian and Duarte, 2020; Adrian et al., 2020), and our estimates provide target moments for such models. A further avenue for future work is the relation between the shape of the conditional distribution of consumption growth and the shape of option-implied expectations for equity returns. Strong downside risks in consumption growth should coincide with downside risks for equities. While the theory makes clear predictions (Barro and Liao, 2021), previous work has lacked an estimate of the conditional distribution of consumption growth. Our estimates could fill this gap and allow tests of the consumption-based model across states of the world, beyond the time series and the cross section of returns.

References

Acosta, M., A. Ajello, M. Bauer, F. Loria, and S. Miranda-Agrippino (2025): "Financial Market Effects of FOMC Communication: Evidence from a New Event-Study Database," Working Paper 2025-30, Federal Reserve Bank of San Francisco.

Adrian, T., N. Boyarchenko, and D. Giannone (2019): "Vulnerable growth," *American Economic Review*, 109, 1263–89.

Adrian, T. and F. Duarte (2020): "Financial Vulnerability and Monetary Policy," Working Paper.

Adrian, T., F. Duarte, N. Liang, and P. Zabczyk (2020): "NKV: A New Keynesian Model with Vulnerability," in *AEA Papers and Proceedings*, vol. 110, 470–476.

Adrian, T., F. Grinberg, N. Liang, S. Malik, and J. Yu (2022): "The term structure of growth-at-risk," *American Economic Journal: Macroeconomics*, 14, 283–323.

Adrian, T. and H. Shin (2010): "Chapter 12 - Financial Intermediaries and Monetary Economics," Elsevier, vol. 3 of *Handbook of Monetary Economics*, 601–650.

Ajello, A. and T. Pike (2022): "Getting in all the Cracks: Monetary Policy, Financial Vulnerabilities, and Macro Risk," *Working Paper*.

Ascari, G. and T. Haber (2022): "Non-linearities, state-dependent prices and the transmission mechanism of monetary policy," *The Economic Journal*, 132, 37–57.

Azzalini, A. and A. Capitanio (2003): "Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65, 367–389.

Bachmann, R. and C. Bayer (2014): "Investment dispersion and the business cycle," *American Economic Review*, 104, 1392–1416.

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

——— (2008): "Forecasting economic time series using targeted predictors," *Journal of Econometrics*, 146, 304–317.

Baker, S. R., N. Bloom, and S. J. Davis (2016): "Measuring economic policy uncertainty," *The Quarterly Journal of Economics*, 131, 1593–1636.

Baqaee, D. R. and E. Farhi (2019): "The macroeconomic impact of microeconomic shocks: Beyond Hulten's theorem," *Econometrica*, 87, 1155–1203.

Barnichon, R. and C. Matthes (2018): "Functional approximation of impulse responses," *Journal of Monetary Economics*, 99, 41–55.

Barro, R. J. (2006): "Rare disasters and asset markets in the twentieth century," *The Quarterly Journal of Economics*, 121, 823–866.

Barro, R. J. and G. Y. Liao (2021): "Rare disaster probability and options pricing," *Journal of Financial Economics*, 139, 750–769.

Basu, S. and B. Bundick (2017): "Uncertainty shocks in a model of effective demand," *Econometrica*, 85, 937–958.

Bauer, M. D. and M. Chernov (2024): "Interest Rate Skewness and Biased Beliefs," *The Journal of Finance*, 79, 173–217.

Beason, T. and D. Schreindorfer (2022): "Dissecting the Equity Premium," *Journal of Political Economy*, 130, 2203–2222.

Beaudry, P. and F. Portier (2006): "Stock prices, news, and economic fluctuations,"

American Economic Review, 96, 1293–1307.

Bekaert, G. and M. Hoerova (2014): "The VIX, the variance premium and stock market volatility," *Journal of Econometrics*, 183, 181–192.

Bekaert, G., M. Hoerova, and M. L. Duca (2013): "Risk, uncertainty and monetary policy," *Journal of Monetary Economics*, 60, 771–788.

Belloni, A. and V. Chernozhukov (2011): "L1-penalized quantile regression in high-dimensional sparse models," *The Annals of Statistics*, 39, 82–130.

Ben Zeev, N. and H. Khan (2015): "Investment-specific news shocks and US business cycles," *Journal of Money, Credit and Banking*, 47, 1443–1464.

Bernanke, B. S. (1983): "Irreversibility, uncertainty, and cyclical investment," *The Quarterly Journal of Economics*, 98, 85–106.

Bernanke, B. S., M. Gertler, and S. Gilchrist (1999): "The Financial Accelerator in a Quantitative Business Cycle Framework," 53.

Bernanke, B. S. and K. N. Kuttner (2005): "What explains the stock market's reaction to Federal Reserve policy?" *The Journal of Finance*, 60, 1221–1257.

Bianchi, F., S. C. Ludvigson, and S. Ma (2022): "Belief distortions and macroeconomic fluctuations," *American Economic Review*, 112, 2269–2315.

Bloom, N. (2009): "The impact of uncertainty shocks," *econometrica*, 77, 623–685.

——— (2014): "Fluctuations in uncertainty," *Journal of Economic Perspectives*, 28, 153–176.

Bloom, N., M. Floetotto, N. Jaimovich, I. Saporta-Eksten, and S. J. Terry (2018): "Really uncertain business cycles," *Econometrica*, 86, 1031–1065.

Boguth, O. and L.-A. Kuehn (2013): "Consumption volatility risk," *The Journal of Finance*, 68, 2589–2615.

Boivin, J. and S. Ng (2006): "Are more data always better for factor analysis?" *Journal of Econometrics*, 132, 169–194.

Bollerslev, T., G. Tauchen, and H. Zhou (2009): "Expected stock returns and variance risk premia," *The Review of Financial Studies*, 22, 4463–4492.

Bollerslev, T. and V. Todorov (2011a): "Estimation of jump tails," *Econometrica*, 79, 1727–1783.

——— (2011b): "Tails, fears, and risk premia," *The Journal of Finance*, 66, 2165–2211.

Bollerslev, T., V. Todorov, and L. Xu (2015): "Tail risk premia and return predictability," *Journal of Financial Economics*, 118, 113–134.

Borio, C. and H. Zhu (2012): "Capital regulation, risk-taking and monetary policy: a missing link in the transmission mechanism?" *Journal of Financial stability*, 8, 236–251.

Breeden, D. T. (1979): "An intertemporal asset pricing model with stochastic consumption and investment opportunities," *Journal of Financial Economics*, 7, 265–296.

Breeden, D. T., M. R. Gibbons, and R. H. Litzenberger (1989): "Empirical tests of the consumption-oriented CAPM," *The Journal of Finance*, 44, 231–262.

Brunnermeier, M. K. and Y. Sannikov (2014): "A macroeconomic model with a financial sector," *American Economic Review*, 104, 379–421.

Bryzgalova, S., J. Huang, and C. Julliard (2023): "Consumption in Asset Returns," *Working paper*.

Busch, C., D. Domeij, F. Guvenen, and R. Madera (2022): "Skewed idiosyncratic income risk over the business cycle: Sources and insurance," *American Economic Journal: Macroeconomics*, 14, 207–42.

Campbell, J. Y. and J. Ammer (1993): "What moves the stock and bond markets? A variance decomposition for long-term asset returns," *The Journal of Finance*, 48, 3–37.

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

Chernozhukov, V., I. Fernández-Val, and A. Galichon (2010): "Quantile and probability curves without crossing," *Econometrica*, 78, 1093–1125.

Christelis, D., D. Georgarakos, T. Jappelli, and M. van Rooij (2020): "Consumption Uncertainty and Precautionary Saving," *Review of Economics and Statistics*, 102, 148–161.

Clark, T. E. and M. W. McCracken (2005): "Evaluating direct multistep forecasts," *Econometric Reviews*, 24, 369–404.

Cochrane, J. (2009): *Asset pricing: Revised edition*, Princeton University Press.

Cochrane, J. H. (2017): "Macro-Finance," *Review of Finance*, 21, 945–985.

Daniel, K., L. Garlappi, and K. Xiao (2021): "Monetary Policy and Reaching for Income," *The Journal of Finance*, 76, 1145–1193.

Drechsler, I., A. Savov, and P. Schnabl (2018): "A model of monetary policy and risk premia," *The Journal of Finance*, 73, 317–373.

Duprey, T. and A. Ueberfeldt (2020): "Managing GDP tail risk," *Bank of Canada Staff Working Paper No. 2020-03*.

Fama, E. F. and K. R. French (1993): "Common risk factors in the returns on stocks and bonds," *Journal of Financial Economics*, 33, 3–56.

Fama, E. F. and J. D. MacBeth (1973): "Risk, return, and equilibrium: Empirical tests," *Journal of Political Economy*, 81, 607–636.

Favara, G., S. Gilchrist, K. F. Lewis, and E. Zakrajsek (2016): "Recession risk and the excess bond premium," *FEDS notes No. 2016-04-08*.

Fernald, J. (2014): "A quarterly, utilization-adjusted series on total factor productivity," *Federal Reserve Bank of San Francisco Working Paper 2012-09*.

Ferreira, T. R. (2018): "Stock market cross-sectional skewness and business cycle fluctuations," *International Finance Discussion Papers 1223*.

Gabaix, X. (2012): "Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance," *The Quarterly Journal of Economics*, 127, 645–700.

Gertler, M. and P. Karadi (2015): "Monetary policy surprises, credit costs, and economic activity," *American Economic Journal: Macroeconomics*, 7, 44–76.

Giannone, D., M. Lenza, and G. E. Primiceri (2021): "Economic predictions with big data: The illusion of sparsity," *Econometrica*, 89, 2409–2437.

Giglio, S., B. Kelly, and S. Pruitt (2016): "Systemic risk and the macroeconomy: An empirical evaluation," *Journal of Financial Economics*, 119, 457–471.

Gilchrist, S. and E. Zakrajšek (2012): "Credit spreads and business cycle fluctuations," *American Economic Review*, 102, 1692–1720.

Gourio, F. (2012): "Disaster risk and business cycles," *American Economic Review*, 102, 2734–2766.

Guiso, L. and T. Jappelli (2024): "Anatomy of Consumption Risk," CSEF Working Paper 732, Centre for Studies in Economics and Finance.

Gürkaynak, R. S., B. P. Sack, and E. T. Swanson (2005): "Do actions speak louder than words? The response of asset prices to monetary policy actions and statements," *International Journal of Central Banking*, 1, 55–93.

Hamilton, J. D. (2003): "What is an oil shock?" *Journal of econometrics*, 113, 363–398.
Hanson, S. G. and J. C. Stein (2015): "Monetary Policy and Long-Term Real Rates," *Journal of Financial Economics*, 115, 429–448.
Hattori, M., A. Schrimpf, and V. Sushko (2016): "The response of tail risk perceptions to unconventional monetary policy," *American Economic Journal: Macroeconomics*, 8, 111–136.
Hodrick, R. J. (1992): "Dividend yields and expected stock returns: Alternative procedures for inference and measurement," *The Review of Financial Studies*, 5, 357–386.
Hyland, C. C. (2026): "Taming Black Swans: Monetary Policy in the Presence of Tail Risks," Working Paper, SSRN 6240078.
Jarociński, M. and P. Karadi (2020): "Deconstructing monetary policy surprises: The role of information shocks," *American Economic Journal: Macroeconomics*, 12, 1–43.
Jordà, Ò. (2005): "Estimation and inference of impulse responses by local projections," *American Economic Review*, 95, 161–182.
Jurado, K., S. C. Ludvigson, and S. Ng (2015): "Measuring uncertainty," *American Economic Review*, 105, 1177–1216.
Kandel, S. and R. F. Stambaugh (1990): "Expectations and volatility of consumption and asset returns," *The Review of Financial Studies*, 3, 207–232.
Känzig, D. R. (2021): "The Macroeconomic Effects of Oil Supply News: Evidence from OPEC Announcements," *American Economic Review*, 111, 1092–1125.
Kelly, B. and H. Jiang (2014): "Tail risk and asset prices," *The Review of Financial Studies*, 27, 2841–2871.
Kilian, L. (2009): "Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market," *The American Economic Review*, 99, 1053–1069.
Kiyotaki, N. and J. Moore (1997): "Credit Cycles," *Journal of Political Economy*, 105, 211–248.
Knox, B., J. M. Londono, M. Samadi, and A. Vissing-Jørgensen (2024): "Equity Premium Events," *Working Paper*.
Knox, B. and A. Vissing-Jørgensen (2022): "A stock return decomposition using observables," *FEDS Working Paper No. 2022-14*.
Koenker, R. and J. A. Machado (1999): "Goodness of Fit and Related Inference Processes for Quantile Regression," *Journal of the American Statistical Association*, 94, 1296–1310.
Kroencke, T. A. (2017): "Asset pricing without garbage," *The Journal of Finance*, 72, 47–98.
Loria, F., C. Matthes, and D. Zhang (2023): "Assessing macroeconomic tail risk," *Working paper*.
Lucas, R. E. (1978): "Asset prices in an exchange economy," *Econometrica*, 46, 1429–1445.
Ludvigson, S. C., S. Ma, and S. Ng (2021): "Uncertainty and business cycles: exogenous impulse or endogenous response?" *American Economic Journal: Macroeconomics*, 13, 369–410.
Mankiw, N. G. and M. D. Shapiro (1986): "RISK AND RETURN: CONSUMPTION BETA VERSUS MARKET BETA." *Review of Economics & Statistics*, 68.
Martin, I. W. and S. Nagel (2022): "Market efficiency in the age of big data," *Journal of Financial Economics*, 145, 154–177.
McCracken, M. W. and S. Ng (2016): "FRED-MD: A Monthly Database for Macroeconomic Research," *Journal of Business & Economic Statistics*, 34, 574–589.

Miranda-Agrippino, S. and G. Ricco (2021): "The transmission of monetary policy shocks," *American Economic Journal: Macroeconomics*, 13, 74–107.

Nagel, S. and Z. Xu (2024): "Movements in Yields, not the Equity Premium: Bernanke-Kuttner Redux," *Working paper*.

Nakamura, E. and J. Steinsson (2018): "High-frequency identification of monetary non-neutrality: the information effect," *The Quarterly Journal of Economics*, 133, 1283–1330.

Nakamura, E., J. Steinsson, R. Barro, and J. Ursúa (2013): "Crises and recoveries in an empirical model of consumption disasters," *American Economic Journal: Macroeconomics*, 5, 35–74.

Parker, J. A. and C. Julliard (2005): "Consumption risk and the cross section of expected returns," *Journal of Political Economy*, 113, 185–222.

Plagborg-Møller, M., L. Reichlin, G. Ricco, and T. Hasenzagl (2020): "When is Growth at risk?" *Brookings Papers on Economic Activity, Spring 2020*.

Rietz, T. A. (1988): "The equity risk premium: A solution," *Journal of Monetary Economics*, 22, 117–131.

Salgado, S., F. Guvenen, and N. Bloom (2023): "Skewed business cycles," *Working paper*.

Savov, A. (2011): "Asset pricing with garbage," *The Journal of Finance*, 66, 177–201.

Schularick, M. and A. M. Taylor (2012): "Credit booms gone bust: monetary policy, leverage cycles, and financial crises, 1870–2008," *American Economic Review*, 102, 1029–1061.

Shanken, J. (1992): "On the estimation of beta-pricing models," *The review of financial studies*, 5, 1–33.

Tenreyro, S. and G. Thwaites (2016): "Pushing on a string: US monetary policy is less powerful in recessions," *American Economic Journal: Macroeconomics*, 8, 43–74.

Wachter, J. A. (2013): "Can time-varying risk of rare disasters explain aggregate stock market volatility?" *The Journal of Finance*, 68, 987–1035.

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

Wu, J. C. and F. D. Xia (2016): "Measuring the macroeconomic impact of monetary policy at the zero lower bound," *Journal of Money, Credit and Banking*, 48, 253–291.

A Consumption forecasting comparison

We compare twenty alternative models for forecasting year-on-year aggregate real consumption growth excluding durables. Each model condenses a high-dimensional predictor panel into factor estimates and uses those factors as right-hand-side variables in the quantile regressions of equation (3). The predictor panel is the February 2026 FRED-MD set of McCracken and Ng (2016), comprising 126 monthly macroeconomic and financial variables. The aligned consumption and factor sample runs from January 1960 to September 2025, and forecasts are generated recursively: the first uses information through January 1970 to predict year-on-year consumption growth in January 1971, and the estimation window then expands one month at a time, with all factors re-estimated at each step.

The first eight models use the first one through eight principal components of the FRED-MD panel as predictors and are labeled PC1 through PC8.[13] Principal components describe the most variation in the predictor panel but do not target the predictive content of individual predictors for the dependent variable. The next model, PQR, instead applies the partial-quantile-regression construction of Giglio et al. (2016) to the full FRED-MD sample to build a single targeted factor. The PQR-PC variant builds the same targeted factor from the eight principal components rather than from the raw FRED-MD panel, in case feeding so many predictors directly into the PQR is too noisy.

Many FRED-MD series are highly correlated, which can cause factor estimation to overweight certain features of the data and induce cross-correlation in the idiosyncratic errors. Pre-selecting a smaller set of predictors before factor estimation can mitigate both issues. The next eight models, labeled PC1$*$ through PC8$*$, therefore pre-select 40 predictors via ℓ_1-penalized quantile regression as in equation (4) before extracting one through eight principal components.[14] The final two models combine pre-selection with partial quantile regression: PQR$*$ extracts a single targeted factor from the pre-selected 40 predictors, and PQR-PC$*$ extracts a single targeted factor from the eight pre-selected principal components in PC8$*$.

Table 10 reports the out-of-sample R^1 averaged across the targeted quantiles $\Theta = \{0.05, 0.25, 0.5, 0.75, 0.95\}$.[15] The first column covers the full 1970–2025 sample, the second

[13] The Bai and Ng (2002) PC_{p2} criterion identifies eight significant factors in the vast majority of periods. Before estimating factors we apply the stationarity transformations recommended in McCracken and Ng (2016), losing the first two observations because some series are second-differenced, and we remove outliers and series with missing values, which drops up to five additional variables.

[14] Boivin and Ng (2006) find that factors built from 40 pre-screened variables can outperform factors estimated from over 100 macroeconomic time series.

[15] The baseline-model entry differs slightly from the corresponding row in Table 1 of the main text because the optimal penalty level λ in (4) is selected with random tie-breaking.

Table 10: Forecasting model comparison (R^1 in %)

	Full sample	Shorter sample	Lower quantiles
PC1	7.6	7.1	9.3
PC2	-1.7	-4.7	-5.5
PC3	-10.1	-15.2	-19.3
PC4	-7.1	-10.7	-15.7
PC5	-6.7	-10.9	-14.8
PC6	-7.1	-11.5	-13.3
PC7	-6.3	-11.2	-12.4
PC8	-5.4	-10.1	-10.8
PQR	11.2	6.4	7.7
PQR-PC	2.3	-1.2	1.5
PC1∗	6.7	7.0	8.4
PC2∗	-2.1	-5.9	-5.6
PC3∗	-5.0	-9.4	-10.4
PC4∗	-5.3	-8.9	-10.8
PC5∗	-5.2	-9.2	-10.2
PC6∗	-4.7	-8.2	-9.7
PC7∗	-3.5	-7.2	-8.5
PC8∗	-3.1	-6.7	-7.1
PQR∗	**11.4**	**8.0**	**9.5**
PQR-PC∗	0.1	-2.9	-2.0

The value in the table is the average R^1 across the targeted quantiles, in percent. The highest value in each column is in boldface. The full sample is 01/1970–09/2025. The shorter sample starts in 01/1980 to exclude the volatile 1970s. The third column reports the average fit over the 5, 25, and 50% quantiles.

restricts attention to the post-1979 sample to exclude the volatile 1970s, and the third averages only over the lower-tail quantiles $\{0.05, 0.25, 0.5\}$. PQR∗, which combines pre-selection with partial quantile regression, has the highest average fit in all three columns.

B Calibration of the conditional consumption growth distribution

A direct calibration check compares the recursive quantile forecasts with realized future consumption growth. For each forecast origin t, we pair the predicted quantiles with the realized year-on-year consumption growth observed at $t + 12$ and compute, for each

quantile index θ, the empirical coverage rate

$$\frac{1}{T}\sum_{t}\mathbb{I}\{c_{t,t+12} \leq \widehat{Q}_{\theta,t}\}.$$

Under exact unconditional calibration, the empirical frequencies should equal the nominal probabilities θ.

Table 11 reports the empirical coverage rates for the same 12-month-horizon recursive forecasts used in the main analysis. The frequencies do not line up exactly with the nominal probabilities, but they are informative about the scale of the forecast distribution. The lower tail is in roughly the right range: the 10% quantile has empirical coverage of 14.2%, and realizations fall below the 5% quantile 11.3% of the time. The model therefore does not deliver exact lower-tail coverage, but it identifies a region of the distribution where realized downside outcomes are rare. Coverage is below nominal from the 25% through the 95% quantiles, with the largest deviations around the median and upper-middle: 33.2% at the median, 52.8% at the 75% quantile, and 75.8% at the 90% quantile. Figure 10 shows the same pattern visually.

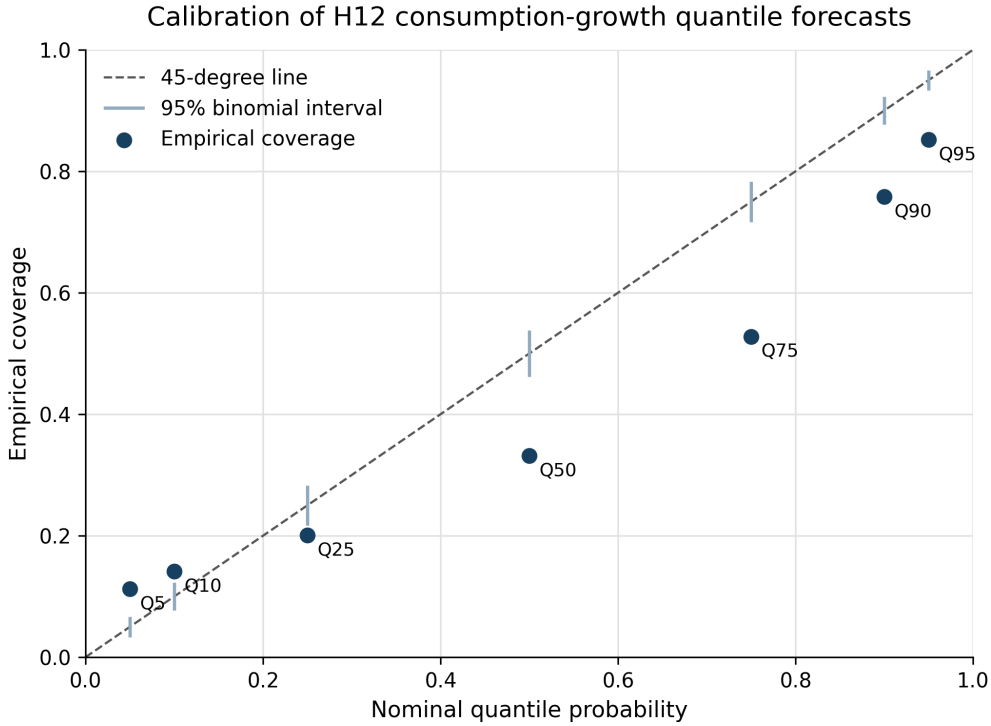
The calibration exercise therefore points to approximate rather than exact probability calibration. This is unsurprising in the present setting: forecast targets overlap at a 12-month horizon, macroeconomic volatility varies over time, and the distributional approximation is constructed from a finite grid of estimated quantiles. The fitted distribution is best read as a disciplined summary of the time-varying shape of expected consumption growth risks rather than as a set of exact probability statements at every percentile. This interpretation aligns with how the distribution is used elsewhere in the paper: the analysis relies on relative features of the estimated distribution, particularly Kelley skewness and related downside risk measures, and evaluates their economic content through out-of-sample forecast performance and return predictability tests.

Table 11: Calibration of recursive consumption-growth quantile forecasts

Quantile	Nominal	Empirical	Error	N	95% binomial CI	Kupiec p
Q5	5.0	11.3	6.3	657	[3.3, 6.7]	< 0.01
Q10	10.0	14.2	4.2	657	[7.7, 12.3]	< 0.01
Q25	25.0	20.1	-4.9	657	[21.7, 28.3]	< 0.01
Q50	50.0	33.2	-16.8	657	[46.2, 53.8]	< 0.01
Q75	75.0	52.8	-22.2	657	[71.7, 78.3]	< 0.01
Q90	90.0	75.8	-14.2	657	[87.7, 92.3]	< 0.01
Q95	95.0	85.2	-9.8	657	[93.3, 96.7]	< 0.01

The table evaluates the H12 recursive forecasts of year-on-year real personal consumption expenditures growth excluding durables. Each row is dated by the forecast origin t, and the realized target is the consumption-growth observation at $t+12$. For each nominal quantile, empirical coverage is the share of realized future consumption-growth observations less than or equal to the model-implied forecast quantile. Error is empirical minus nominal coverage, in percentage points. The 95% confidence interval is the asymptotic binomial interval under correct unconditional coverage; because H12 forecast targets overlap, the interval should be read as a diagnostic benchmark rather than a full small-sample inference procedure. The Kupiec column reports the unconditional-coverage test p-value.

Figure 10: Empirical coverage of recursive consumption growth quantile forecasts



Note: The figure plots nominal quantile probabilities against empirical coverage rates for the H12 recursive consumption growth forecasts. The vertical intervals are asymptotic 95% binomial intervals under correct unconditional coverage. Because forecast targets overlap at a 12-month horizon, the intervals are diagnostic benchmarks rather than full small-sample inference.

C Selected predictors for consumption forecasting

Which predictors are particularly useful in forecasting quantiles of the consumption growth distribution? Table 12 lists the ten most frequently selected predictors from the pre-selection via ℓ_1-penalized quantile regression. For the 5% quantile, the most frequently chosen variables are the inventories-to-sales ratio (ISRATIOx), real personal income (RPI), and the Canadian dollar exchange rate (EXCAUSx). Credit and short-rate spreads also feature prominently through COMPAPFFx and TB3SMFFM, while labor-market indicators such as UNRATE, CES1021000001, and AWOTMAN are common selections. Regional housing starts, especially HOUSTS and HOUSTNE, also appear regularly. The median quantile again loads heavily on ISRATIOx and EXCAUSx, but housing permits, consumer credit, manufacturing hours, and the S&P price-earnings ratio become more prominent. For the 95% quantile, the selected set leans more toward income, exchange rates, credit, and asset-market variables, with RPI, EXCAUSx, CONSPI, DTCOLNVHFNM, and the S&P 500 all appearing among the ten most frequent selections.

Two caveats are in order. First, the ℓ_1 penalty typically selects only one variable from a group of highly correlated regressors and is agnostic about which one (Bai and Ng, 2008, p. 307). Each listed predictor should therefore be read as a representative of a correlated cluster rather than as the unique informative variable. Second, selection frequency provides only indirect evidence on importance, since the final factor weight is determined in the partial-quantile-regression step rather than at the pre-selection stage.

To show how variable selection evolves over time, we group the predictors into the eight FRED-MD blocks of McCracken and Ng (2016): (1) output and income, (2) labor market, (3) housing, (4) consumption, orders, and inventories, (5) money and credit, (6) interest and exchange rates, (7) prices, and (8) stock market. For each out-of-sample month we compute the share of the 40 selected predictors that falls into each block. Figures 11 and 12 report the resulting stacked shares for the 5% and 95% quantiles, complementing the individual-predictor evidence in Table 12 by characterizing the pre-selected information set rather than the final factor weights.

Figure 11 shows that lower-tail forecasts draw heavily on financial conditions, especially interest and exchange rates and money and credit variables. Prices account for a large share in parts of the 1970s, while housing and labor-market variables become more visible in later episodes. The composition moves considerably over time: the downside tail is not forecast by a fixed set of predictors but by a rotating mix of macroeconomic and financial variables. This pattern is consistent with the main results in the paper, in which monetary and financial conditions are especially informative about downside

Table 12: Most frequently selected predictors

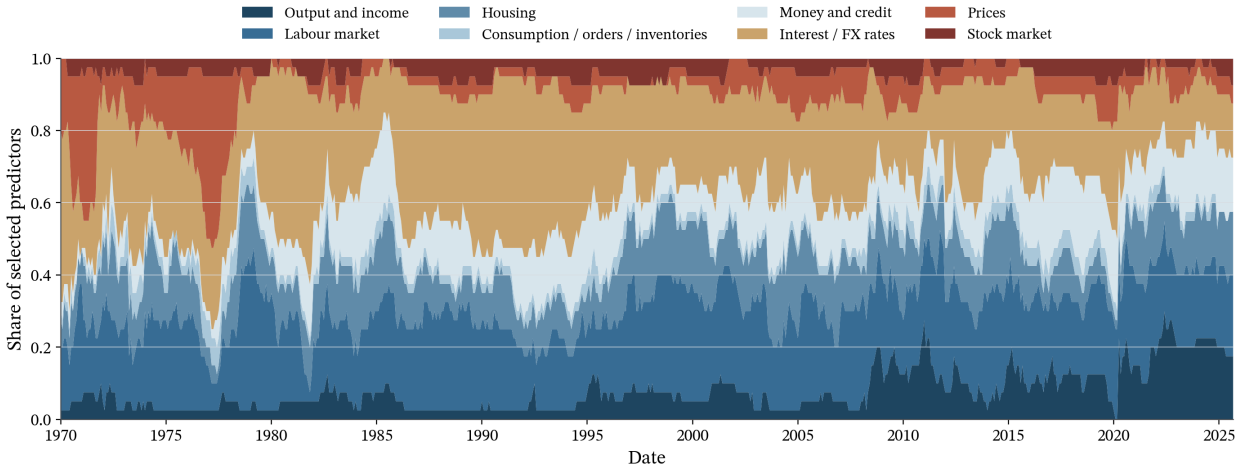
	5% Quantile		Median		95% Quantile	
	Name	Frequency	Name	Frequency	Name	Frequency
1	ISRATIOx	96.0	ISRATIOx	90.9	RPI	89.2
2	RPI	91.0	EXCAUSx	89.2	EXCAUSx	81.8
3	EXCAUSx	88.6	PERMITS	82.4	CONSPI	79.1
4	COMPAPFFx	84.5	CONSPI	82.2	ISRATIOx	78.3
5	TB3SMFFM	80.0	CES1021000001	81.5	DTCOLNVHFNM	76.5
6	UNRATE	74.9	PERMITMW	81.2	EXSZUSx	74.1
7	CES1021000001	72.5	EXUSUKx	80.1	CES1021000001	71.6
8	HOUSTS	69.5	S&P P/E Ratio	77.3	EXJPUSx	71.3
9	AWOTMAN	68.0	AWHMAN	75.8	WPSID62	70.7
10	HOUSTNE	67.9	HWIURATIO	74.9	S&P 500	69.5

Legend: AWHMAN – Average weekly hours: manufacturing. AWOTMAN – Average weekly overtime hours: manufacturing. CES1021000001 – All employees: mining and logging: mining. COMPAPFFx – 3-month commercial paper minus federal funds rate. CONSPI – Nonrevolving consumer credit to personal income. DTCOLNVHFNM – Consumer motor vehicle loans outstanding. EXCAUSx – Canada/US foreign exchange rate. EXJPUSx – Japan/US foreign exchange rate. EXSZUSx – Switzerland/US foreign exchange rate. EXUSUKx – US/UK foreign exchange rate. HOUSTNE – Housing Starts, Northeast. HOUSTS – Housing Starts, South. HWIURATIO – Help wanted to number of unemployed. ISRATIOx – Total business: inventories to sales. PERMITMW – New private housing permits, Midwest (SAAR). PERMITS – New private housing permits, South (SAAR). RPI – Real personal income. S&P 500 – S&P composite index level. S&P P/E Ratio – P/E ratio on S&P composite index. TB3SMFFM – 3-month treasury yield minus federal funds rate. UNRATE – Civilian unemployment rate. WPSID62 – PPI Crude Materials.

consumption risk while real-activity indicators still matter for the location and timing of tail events.

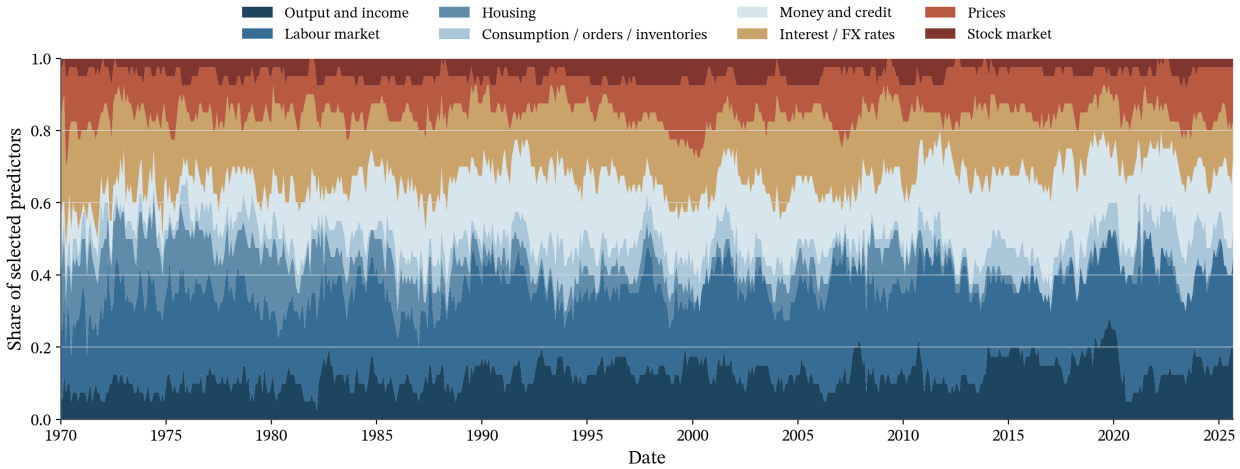
Figure 12 displays a more stable mix for the upper tail. Output and income, labor-market indicators, money and credit, and interest and exchange rates each contribute persistently, while housing is less dominant than in the lower-tail figure. Consumption, orders, and inventories represent a small group share in both figures, even though the inventories-to-sales ratio is one of the most frequently selected individual predictors in Table 12. Taken together, the figures support the use of a broad macro-financial panel: the consumption growth distribution is forecast using information from many parts of the economy, and the relevant blocks differ across the lower and upper tails.

Figure 11: Predictor-group composition for the 5% consumption growth quantile



Note: For each out-of-sample month, the stacked areas show the share of the 40 predictors selected for the 12-month-ahead 5% quantile of real PCE consumption growth excluding durables that belongs to each variable group. The group definitions follow the appendix of McCracken and Ng (2016).

Figure 12: Predictor-group composition for the 95% consumption growth quantile



Note: For each out-of-sample month, the stacked areas show the share of the 40 predictors selected for the 12-month-ahead 95% quantile of real PCE consumption growth excluding durables that belongs to each variable group. The group definitions follow the appendix of McCracken and Ng (2016).

Table 13: Consumption left-tail risk and NY Fed GDP forecast-error risk

Panel A. Summary statistics Measure	Start	End	*N*	Mean	SD	Min	Max
Consumption left-tail risk (-KSK)	1989-01	2025-09	441	0.058	0.312	-0.807	0.789
NY Fed GDP forecast-error left-tail risk (-KSK)	1989-01	2026-01	445	0.046	0.081	-0.286	0.275
NY Fed GDP forecast-error P10 risk (-P10)	1989-01	2026-01	445	2.409	1.118	1.245	10.434

Panel B. Correlations Pair	Full	*N*	Recession	*N*	Expansion	*N*
Consumption KSK vs GDP FE P10	-0.129	441	-0.067	36	-0.011	399
Consumption KSK vs GDP FE KSK	-0.196	441	0.035	36	-0.236	399
Left-tail risk: consumption vs GDP FE KSK	-0.196	441	0.035	36	-0.236	399
Consumption left-tail risk vs GDP FE P10 risk	-0.129	441	-0.067	36	-0.011	399

Note: Consumption left-tail risk is −KSK from the conditional consumption growth distribution. GDP forecast-error left-tail risk is −KSK computed from the New York Fed GDP forecast-error percentiles. GDP forecast-error P10 risk is the negative of the GDP forecast-error 10th percentile. Higher values indicate worse downside risk. Correlations are reported for the full sample and separately by NBER recession status when the recession indicator is available.

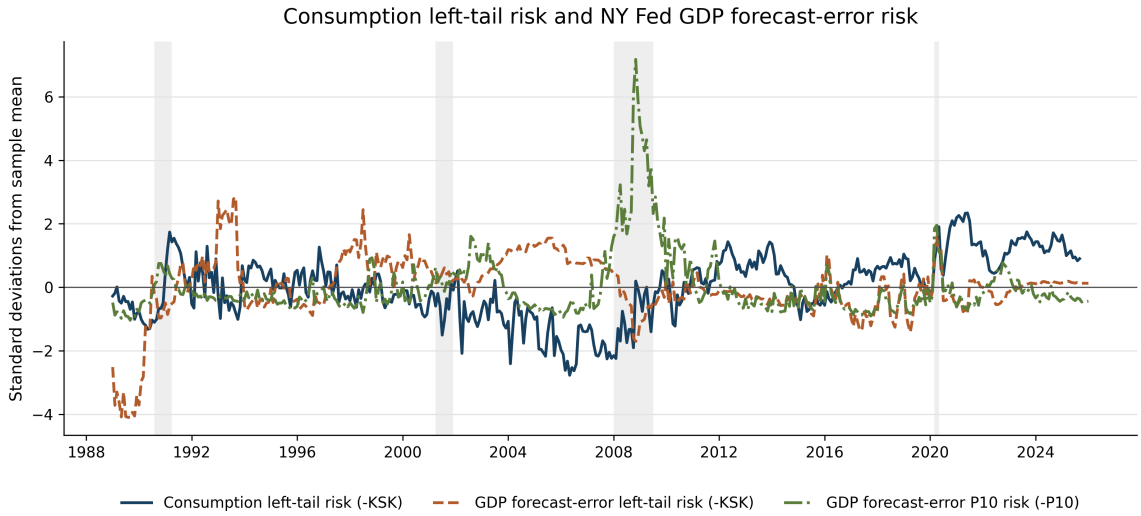
D NY Fed GDP forecast-error risk benchmark

We benchmark consumption left-tail risk against the New York Fed Outlook-at-Risk series for real GDP growth, a public implementation of the Adrian et al. (2019) growth-at-risk methodology. The series reports conditional percentiles of the forecast-error distribution for one-year-ahead real GDP growth. Outcome-level GDP-at-Risk percentiles obtain by adding the Blue Chip Economic Indicators consensus mean back to each forecast-error percentile; because this common location shift leaves Kelley skewness and percentile spreads unchanged, we work with the forecast-error percentiles directly.

Table 13 compares the sign-adjusted risk measures over the merged monthly sample. Consumption left-tail risk is minus consumption Kelley skewness, GDP forecast-error left-tail risk is minus GDP forecast-error Kelley skewness, and GDP forecast-error P10 risk is minus the GDP forecast-error 10th percentile, so higher values always indicate worse downside risk. The merged panel contains 445 monthly observations from January 1989 to January 2026, and the return predictability regressions below use the largest feasible common sample at each return horizon. The two objects are not redundant: consumption left-tail risk has a full-sample correlation of −0.196 with GDP forecast-error KSK risk and −0.129 with GDP forecast-error P10 risk.

Figure 13 plots the standardized series. The two measures move differently across episodes: GDP forecast-error P10 risk spikes sharply around the Great Recession, whereas consumption left-tail risk declines before the crisis and rises more visibly after 2020. The NY Fed GDP forecast-error distribution is a useful external benchmark, but it does not

Figure 13: Consumption left-tail risk and NY Fed GDP forecast-error risk



Notes: Series are sign-adjusted so higher values indicate worse downside risk.
NY Fed values are forecast-error percentiles; outcome-level GDP-at-Risk requires adding the BCEI mean forecast.

Note: The figure plots standardized monthly series. All measures are sign-adjusted so higher values indicate worse downside risk. The shaded regions are NBER recessions. The NY Fed variables are forecast-error risks; outcome-level GDP-at-Risk percentiles require adding the Blue Chip Economic Indicators mean forecast.

summarize the same variation as the consumption risk measure used in the asset pricing tests.

Table 14 reports return predictability horse races. The dependent variable is the cumulative future Fama-French market excess return at horizons $h \in \{1, 3, 6, 12\}$ months. Predictors are standardized so coefficients are comparable across the consumption and GDP measures, and standard errors use the Hodrick (1992) correction. Consumption left-tail risk remains positive and statistically meaningful at all four horizons after controlling for either GDP forecast-error P10 risk or GDP forecast-error KSK risk, while the GDP-benchmark coefficients are imprecise in the horse-race specifications. The comparison supports the interpretation that consumption left-tail risk captures asset pricing variation not subsumed by a public growth-at-risk benchmark.

Table 14: Return predictability horse races with NY Fed GDP forecast-error risk

h	Model	Cons. risk	GDP FE P10 risk	GDP FE KSK risk	R^2	N
1	Consumption only	0.70*** (0.19)			2.6	441
1	GDP FE P10 risk only		0.01 (0.31)		0.0	441
1	GDP FE KSK risk only			-0.01 (0.19)	0.0	441
1	Consumption + GDP FE P10	0.72*** (0.19)	0.10 (0.31)		2.6	441
1	Consumption + GDP FE KSK	0.73*** (0.20)		0.14 (0.21)	2.7	441
3	Consumption only	1.75*** (0.56)			5.4	441
3	GDP FE P10 risk only		-0.05 (0.92)		0.0	441
3	GDP FE KSK risk only			-0.01 (0.55)	0.0	441
3	Consumption + GDP FE P10	1.77*** (0.54)	0.18 (0.91)		5.4	441
3	Consumption + GDP FE KSK	1.81*** (0.59)		0.34 (0.58)	5.6	441
6	Consumption only	2.86*** (1.09)			7.1	439
6	GDP FE P10 risk only		0.30 (1.69)		0.1	439
6	GDP FE KSK risk only			-0.28 (1.08)	0.1	439
6	Consumption + GDP FE P10	2.95*** (1.05)	0.67 (1.67)		7.5	439
6	Consumption + GDP FE KSK	2.92** (1.16)		0.29 (1.15)	7.1	439
12	Consumption only	5.05** (2.23)			10.4	433
12	GDP FE P10 risk only		1.40 (2.82)		0.8	433
12	GDP FE KSK risk only			-0.91 (2.09)	0.3	433
12	Consumption + GDP FE P10	5.31** (2.13)	2.05 (2.75)		12.2	433
12	Consumption + GDP FE KSK	5.07** (2.35)		0.09 (2.21)	10.5	433

Note: The dependent variable is the cumulative future excess return on the Fama-French market portfolio over horizon h. Predictors are standardized and sign-adjusted so higher values indicate worse downside risk. GDP FE denotes the New York Fed GDP forecast-error distribution. Standard errors in parentheses use the Hodrick (1992) correction for overlapping returns. Asterisks denote statistical significance at the 10%, 5%, and 1% levels.

E Monetary-policy shock audit and robustness

The body of Section 4 reports the benchmark monetary-policy response, the state-conditioned paths, and the quantile mechanism. This appendix documents the remaining robustness checks: alternative high-frequency surprise series, alternative sample windows, and alternative definitions of the recent-skewness state indicator. The state-dependent specifications take the form

$$\text{ksk}_{t+h} = \alpha_h + \beta_h \Delta\text{mp}_t + \zeta_h f\left(\Delta\text{mp}_t\right) + \sum_{\ell=1}^{L} \gamma'_{\ell,t}\text{controls}_{t-\ell} + e_{t+h}, \tag{18}$$

where $f(x) = \widehat{\text{ksk}}_t x$ and $\widehat{\text{ksk}}_t = (1/V)\sum_{v=1}^{V} \text{ksk}_{t-v}$ is a moving average of recent Kelley skewness. The benchmark uses $V = 6$ months. The coefficient ζ_h measures whether the response of Kelley skewness to a monetary surprise depends on whether downside risk has recently been high or low.

E.1 Monetary-shock medium-window summaries

The choice of the Jarociński and Karadi (2020) monetary-policy component as the benchmark is itself a modelling decision, and we therefore document how the central result depends on it. The high-frequency event data underlying every alternative surprise series come from the U.S. Monetary Policy Event-Study Database (Acosta et al., 2025). The same convention applies throughout: months without an FOMC announcement, or with no measured surprise, are set to zero, and observations outside a shock's available sample are missing.

We summarize each impulse response over horizons six through eighteen months. This medium-horizon window abstracts from short-run announcement noise and avoids the least precise long-horizon estimates, and it removes the temptation to select a single horizon after inspecting the response path. Table 15 reports the natural-sample estimates. The average-response column is the unconditional Kelley skewness response to a one-standard-deviation monetary loosening; the low- and high-KSK columns evaluate the state-dependent response when the six-month moving average of recent Kelley skewness equals −0.4 and +0.4, respectively; and the final column reports the high-minus-low wedge, with negative entries indicating that the estimated easing response is larger in low-KSK/high-downside-risk states. The pattern is consistent with the benchmark JK result: across nearly all surprise series the average response is small, the low-KSK/high-

Table 15: Natural-sample medium-window KSK responses

Shock	Family	Sample	N	Avg. Response	Low KSK	High KSK	High - Low KSK Wedge
MAR	MAR	1994-02–2020-12	299	0.0115	0.0130	0.0124	-0.0006
MAR full AR	MAR	1994-02–2020-12	299	0.0066	0.0070	0.0067	-0.0004
JK MP-PM FF4	JK	1994-02–2025-06	342	0.0016	0.0496	-0.0224	-0.0721
JK MP-PM PC1	JK	1994-02–2025-06	342	-0.0023	0.0223	-0.0159	-0.0383
JK policy PC1	JK	1994-02–2025-06	342	-0.0008	0.0130	-0.0102	-0.0232
Acosta ME	Acosta	1994-02–2025-06	365	-0.0014	0.0075	-0.0021	-0.0096
Acosta statement	Acosta	1994-02–2025-06	365	0.0007	0.0130	-0.0066	-0.0196
GSS target	GSS	1994-02–2025-06	365	-0.0001	0.0202	-0.0211	-0.0412
GSS level PC1	GSS	1994-02–2025-06	365	0.0007	0.0130	-0.0066	-0.0196
JK median	JK	1994-02–2025-06	342	-0.0007	0.0131	-0.0102	-0.0233
GSS path	GSS	1994-02–2025-06	365	0.0019	-0.0142	0.0141	0.0283

downside-risk response is positive, the high-KSK response is negative or close to zero, and the high-minus-low wedge is negative. The Miranda-Agrippino and Ricco (2021) variants (MAR and `mar_full_ar`) deliver positive responses in both states, and we retain MAR as a conservative information-adjusted robustness check rather than as the benchmark.

E.2 Quantile-response diagnostic

Table 16 examines whether the positive Kelley skewness response in low-KSK/high-downside-risk states is accompanied by an improvement in the lower tail of expected consumption growth. We report the medium-window KSK response in the low-KSK/high-downside-risk state alongside the response of the lower-tail gap, $q_{50} - q_{10}$, over the same horizons. A negative $q_{50} - q_{10}$ response means that monetary easing narrows the distance between the median and the 10th percentile, so the lower tail improves relative to the center of the distribution. For the benchmark JK shock and several alternative surprises, the positive KSK response is accompanied by a negative $q_{50} - q_{10}$ response. The state-dependent rise in Kelley skewness is therefore consistent with an asymmetric improvement in the lower tail across surprise definitions, rather than a feature of any one shock construction.

Table 16: Quantile-response mechanism, h=6–18. KSK low is the medium-window Kelley skewness response in the low-KSK/high-downside-risk state; $q_{50} - q_{10}$ is the response of the median-minus-10th-percentile gap in that same state.

Shock	Model	Sample	KSK low	$q_{50} - q_{10}$
JK MP-PM FF4	Baseline	1994-02–2025-06	0.0496	-0.0749
JK MP-PM FF4	KSK lags	1994-02–2025-06	0.0496	-0.0599
MAR	Baseline	1994-02–2020-12	0.0130	-0.0107
MAR	KSK lags	1994-02–2020-12	0.0130	-0.0382
GSS target	Baseline	1994-02–2025-06	0.0202	-0.0094
GSS target	KSK lags	1994-02–2025-06	0.0202	0.0222
Acosta ME	Baseline	1994-02–2025-06	0.0075	0.0320
Acosta ME	KSK lags	1994-02–2025-06	0.0075	0.0060

Note: All entries are averages over horizons six through eighteen months. Negative values in the $q_{50} - q_{10}$ column indicate that monetary easing narrows the left-tail gap, so the 10th percentile improves relative to the median.