The Journal of FINANCE
The Journal of THE AMERICAN FINANCE ASSOCIATION

THE JOURNAL OF FINANCE • VOL. LXXVII, NO. 2 • APRIL 2022

Volatility Expectations and Returns

LARS A. LOCHSTOER and TYLER MUIR

ABSTRACT

We provide evidence that agents have slow-moving beliefs about stock market volatility that lead to initial underreaction to volatility shocks followed by delayed overreaction. These dynamics are mirrored in the VIX and variance risk premiums, which reflect investor expectations about volatility, and are also supported in both surveys and firm-level option prices. We embed these expectations into an asset pricing model and find that the model can account for a number of stylized facts about market returns and return volatility that are difficult to reconcile, including a weak or even negative risk-return trade-off.

AGENTS' PERCEPTIONS OF RISK PLAY a critical role in asset pricing models. However, a long literature finds that the empirical risk-return trade-off is weak at best (Glosten, Jagannathan, and Runkle (1993)), despite this trade-off being strong in leading asset pricing models (Moreira and Muir (2017)).[1] This paper proposes a model in which a representative agent has biased, slow-moving expectations about volatility. We show that these expectations help explain a weak relation between risk and returns. We discipline expectations about volatility in the model in three ways: we microfound beliefs based on sticky and extrapolative expectations (which have been shown to be present in many other contexts), we use survey data to directly assess agents' expectations of

Lars A. Lochstoer is with UCLA Anderson School of Management. Tyler Muir is with UCLA Anderson School of Management and NBER. We thank seminar participants at NBER Behavioral Finance, Virtual Finance Seminar, UCLA, UT Austin, Tilburg, Toronto, Insead, Norwegian School of Economics, Maastricht, Nova, Amsterdam, and Stanford, as well as Andy Atkeson, Nick Barberis, John Campbell, Ing-Haw Cheng, Mike Chernov, Itamar Dreschler, Andrea Eisfeldt, Stefano Giglio, Valentin Haddad, Barney Hartman-Glaser, Hanno Lustig, Alan Moreira, Stefan Nagel (Editor), Stavros Panageas, Andrei Shleifer, Kelly Shue, and two anonymous referees for comments. We thank James O'Neill and Alvaro Boitier for excellent research assistance. We thank Ing-Haw Cheng, Stefano Giglio, Ian Dew-Becker, and Travis Johnson for making data available. The authors have nothing to disclose.

Correspondence: Tyler Muir, UCLA Anderson School of Management and NBER; e-mail: tyler.muir@anderson.ucla.edu.

[1] Moreira and Muir (2017) show that the basic risk-return relation is strong in calibrations of leading equilibrium asset pricing models (including the habit model of Campbell and Cochrane (1999), long-run risk model of Bansal and Yaron (2004), time-varying disasters model of Wachter (2013), and intermediary-based models as in He and Krishnamurthy (2013)). Martin (2016) argues that this relationship holds in a wide class of models if σ_t^2 is replaced by risk-neutral variance, which we will consider empirically as well.

DOI: 10.1111/jofi.13120

volatility, and we study prices and returns of volatility-dependent claims (VIX futures, variance swaps, and straddles), which help assess whether mistakes in beliefs are present in financial market prices.

Expectations about volatility appear to initially underreact to news about volatility followed by a delayed overreaction. This pattern matches direct empirical facts about volatility-dependent claims and the variance risk premium. Cheng (2018) finds that prices of VIX futures do not respond strongly enough to changes in volatility, so that increases in volatility negatively predict the premium on a short position in VIX futures. We extend these results to variance swaps and straddle returns over a longer sample. Claims that provide insurance against future volatility, which are unconditionally expensive, initially appear "too cheap" after volatility rises but appear expensive later on. This is hard to square with a rational risk premium because these claims are typically riskier after volatility rises (Cheng (2018)), suggesting that the risk premium should go up rather than down.

We next embed these beliefs into an otherwise standard (Epstein and Zin (1989)) equilibrium model with stochastic volatility and argue that we can explain many empirical facts about stock market volatility, the VIX, the variance risk premium, and stock returns. First, the model can generate a weak—and potentially even negative—conditional risk-return trade-off. When volatility rises, agents only partially react by requiring a higher expected return on stocks, which pushes current stock prices down. This matches the negative correlation between realized returns and volatility innovations (French, Schwert, and Stambaugh (1987)) and is consistent with a discount rate effect from volatility shocks. Given the initial underreaction, however, agents on average update positively about volatility next period even without additional news. This effect can push next period prices down further, and make it appear as though the initial increase in volatility is associated with a future decline in the observed equity risk premium in the short term. Through this channel, the model can simultaneously match the strong negative correlation between realized stock returns and contemporaneous innovations in volatility while also generating a weak risk-return trade-off. The latter leads to volatility timing strategies that generate positive alpha (Moreira and Muir (2017)). The price decline is hump-shaped and prices take longer to recover than in the rational benchmark, implying that objective equity risk premiums are high well after the volatility shock subsides. We provide evidence consistent with this view, as in Brandt and Kang (2004).

The variance risk premium—defined as VIX^2 minus an objective forecast of variance—also features underreaction and delayed overreaction through the beliefs channel as variance beliefs are reflected in market-implied volatility (VIX). Because the mistake in beliefs shows up in both volatility claims and equity claims in the same direction, the observed variance risk premium will positively forecast stock returns (Bollerslev, Tauchen, and Zhou (2009), Drechsler and Yaron (2011)). However, compared to rational models that link equity and variance risk premiums (Bollerslev, Tauchen, and Zhou (2009)), our model can account for the otherwise puzzling evidence that while the variance risk premium positively forecasts stock returns, neither the VIX nor realized variance is an individually strong forecaster of returns.

We microfound the beliefs in our model through sticky expectations and extrapolation, both of which have been documented in many other contexts (e.g., Coibion and Gorodnichenko (2015), Mankiw and Reis (2002), Bordalo et al. (2020), Landier, Ma, and Thesmar (2019)).[2] These forces combine to generate initial underreaction and delayed overreaction to volatility news. Intuitively, sticky expectations lead to underreaction in beliefs particularly at short horizons, while overextrapolation leads to overreaction. While these forces are prevalent in other work and lead to a convenient and tractable belief process, we do not take a strong stand on these microfoundations—other behavioral factors that lead to initial underreaction and delayed overreaction may also generate similar dynamics (Barberis, Shleifer, and Vishny (1998), Daniel, Hirshleifer, and Subrahmanyam (1998), Hong and Stein (1999)).[3]

We use survey data on volatility and uncertainty about stock returns from the Graham and Harvey CFO survey as well as the Shiller survey and document that the surveys exhibit slow-moving expectations as in our model. In particular, respondents are asked about the 90^{th} and 10^{th} percentiles of stock returns over the next year, which we use to infer beliefs about volatility. We regress survey expectations about volatility on past volatility realizations and show that expectations look like a weighted average of past volatility realizations as our model predicts, whereas optimal forecasts mainly load on current volatility only. Thus, the survey expectations appear sticky or slow-moving. The evidence of slow updating toward volatility is consistent with the dynamics of CFO learning found in Boutros et al. (2019). We also use longer horizon survey evidence based on a 10-year horizon, which allows us to assess agents views on the persistence of volatility. Similar to Landier, Ma, and Thesmar (2019), we find that agents use a persistence parameter that is too large, consistent with overextrapolation. This can generate overreaction in long-run claims on volatility in the spirit of the findings in Giglio and Kelly (2017) and Stein (1989).

The sharpest place to identify such beliefs in financial market prices is in volatility claims and option markets. While survey evidence is useful, a concern is that biases in surveys may not show up in prices (e.g., if respondents do not actively trade or if other rational agents trade sufficiently to eliminate respondents' impact on prices). We follow and extend (Cheng (2018)) and show that the mistakes in the model show up in prices and in predictable variation in the returns of variance swaps, VIX futures, and straddles. We also document results at the firm level, where we show that implied volatility from firm-level

[2] See Payzan-LeNestour, Pradier, and Putniņš (2018) for related work on expectations of volatility based on past volatility. Bouchaud et al. (2019) and Jiang, Krishnamurthy, and Lustig (2018) apply sticky expectations to the profitability anomaly and exchange rates, respectively.

[3] Our model is related to other models of extrapolation from past data or experience effects, including Barberis et al. (2015), Collin-Dufresne, Johannes, and Lochstoer (2016), Nagel and Xu (2019), Malmendier and Nagel (2011), Greenwood and Shleifer (2014), and Glaeser and Nathanson (2017).

options does not react strongly enough to recent changes in volatility, leading to underreaction and a lower variance risk premium following increases in volatility (see Poteshman (2001) for related work). This is true even when we include time fixed effects that control for aggregate movements in firm volatility, which makes a risk-based explanation even more difficult. The firm-level analysis provides further support for our story of initial underreaction and also provides robustness to our main empirical results, which rely on aggregate market data and hence a relatively smaller sample.[4]

We use the survey data and evidence from option markets to calibrate our model. The calibrated model well matches the main quantitative stylized facts in the literature on the relation between volatility, stock returns, and the variance risk premium, which we extend to a more recent sample. Most importantly, we show that slow-moving expectations about volatility are key to matching these dynamics—the nested rational version of the model fails to account for the evidence. We come to similar conclusions for the rational model of Bollerslev, Tauchen, and Zhou (2009). A natural concern is that biases in beliefs will lead to excessive trading profits for a rational investor in the model. The mistakes in the calibrated model are modest: agents' beliefs about volatility in the model have a correlation of about 0.9 with an objective forecast. In an extension, we show that Sharpe ratio gains for a rational investor who trades on these mistakes are similar to other anomalies in the literature. We also explicitly consider learning in our objective volatility forecasts where we construct estimates of variance only using data available to the agents at the time.

Our model is parsimonious, stylized, and tractable, and our main point is that this simple change in volatility expectations can help match key features of the data. Section I provides a microfoundation for the expectations in this paper and brings in survey data about volatility. Section II presents the model, while Section III compares the model to the stylized facts in the literature on stock returns, volatility, the VIX, and the variance risk premium. Section IV provides additional evidence, discusses shortcomings of the model, and considers alternative explanations. Finally, Section V concludes.

I. Belief Microfoundations and Survey Evidence

A. A Simple Microfounded Model of Variance Belief Formation

We provide a simple model of investor expectations of stock market variance that features both sticky expectations and overextrapolation. We show that both features are present in survey data on variance expectations, consistent with a burgeoning literature that estimates beliefs from various surveys and experiments.[5]

[4] Rachwalski and Wen (2016) also provide firm-level evidence on the risk-return trade-off: "Stocks with increases in idiosyncratic risk tend to earn low subsequent returns for a few months. However, high idiosyncratic risk stocks eventually earn persistently high returns."

[5] For example, Coibion and Gorodnichenko (2015) find that average survey expectations display dynamics consistent with the sticky information model of Mankiw and Reis (2002). At the same

As in Mankiw and Reis (2002), a fraction $1-\tau$ of agents update beliefs at each point in time. Denote their time $t-j$ forecast of realized variance (rv) at time $t+1$ by $F_{t-j}[rv_{t+1}]$. We denote the average, or aggregate, expectation of next-period variance across investors by $E_t^S[rv_{t+1}]$, where we use the S superscript to emphasize that these are subjective expectations. Following Mankiw and Reis (2002), the average forecast ($E_t^S[rv_{t+1}]$) is then

$$E_t^S[rv_{t+1}] = (1-\tau)\sum_{j=0}^{\infty}\tau^j F_{t-j}[rv_{t+1}], \tag{1}$$

where $(1-\tau)\tau^j$ is the proportion of agents who last revised their expectation at time $t-j$.[6]

Further, agents believe that realized variance follows an AR(1) process with autocorrelation coefficient $\tilde{\rho}$. Specifically, $F_{t-j}[rv_{t+1}] = \bar{v} + \tilde{\rho}^{j+1}(rv_{t-j}-\bar{v})$, where $\bar{v}$ is the unconditional mean. If $\tilde{\rho}$ is too high relative to the true persistence of expected variance, agents overextrapolate, consistent with the findings in Landier, Ma, and Thesmar (2019). This gives a simple model that relates expectations to lagged realized variance,

$$E_t^S[rv_{t+1}] = \bar{v} + (1-\tau)\tilde{\rho}\sum_{j=0}^{\infty}\tau^j\tilde{\rho}^j(rv_{t-j}-\bar{v}). \tag{2}$$

This form of belief formation is also used in Brooks, Katz, and Lustig (2018). A higher value of τ implies more information stickiness, which leads to initial underreaction in aggregate beliefs as beliefs load too heavily on past lags of realized variance. A higher value of $\tilde{\rho}$ leads to overreaction and too persistent beliefs. This is particularly clear if we consider agents' longer run expectations about variance, as the expected variance k periods from today is

$$E_t^S[rv_{t+k}] = \bar{v}(1-\tilde{\rho}^{k-1}) + \tilde{\rho}^{k-1}E_t^S[rv_{t+1}]. \tag{3}$$

Because stickiness generates initial underreaction and overextrapolation generates overreaction, particularly at longer horizons, the combination of the two effects can lead to initial underreaction followed by delayed overreaction. This simple model provides a useful way to interpret our empirical estimates and gives a microfoundation for the belief dynamics we use in our main model and that we uncover in survey beliefs. We note, however, that belief patterns that generate initial underreaction and delayed overreaction can be generated assuming other investor biases, as in the previous literature (Daniel, Hirshleifer, and Subrahmanyam (1998), Barberis, Shleifer, and Vishny (1998), Hong

time, studies on individual agents' expectations (e.g., Landier, Ma, and Thesmar (2019)) find that agents overextrapolate in the sense that they believe that shocks have a more persistent effect than they truly do.

[6] To see this consider period $t-j$, where a fraction $(1-\tau)$ of agents update at this point in time. Each period going forward a fraction τ of these agents do not update. Hence, at time t (j periods later), the proportion of agents who last updated their expectation at $t-j$ is $(1-\tau)\tau^j$.

and Stein (1999), Fuster, Hebert, and Laibson (2011)). The key feature we emphasize is that agents use too many lags of variance when forming expectations about future variance.

B. Survey Data

Surveys allow us to evaluate the main mechanism in our paper using direct data on variance expectations. Our main source is the Graham and Harvey survey of CFOs, which is quarterly from 2001 to 2018. The survey asks respondents for a mean forecast for the stock market return over the next year as well as 10^{th} and 90^{th} percentiles. We construct the 90^{th} minus 10^{th} percentile as a measure of volatility or uncertainty and square this number to get a measure of expected variance. While this measure has limitations, it captures how spread out agents view the return distribution and, under a Normal distribution, is proportional to agents' expectations about variance.

B.1. Stickiness of Variance Expectations

We fit survey expectations and realized variance as an exponentially weighted average of past variance as in the microfounded model above. Equation (2) implies that the expected 12-month variance from the survey is related to lagged monthly realized variances according to

$$Survey_t^{(1yr)} = a + b\sum_{j=0}^{J}\phi^j rv_{t-j} + \epsilon_t^s. \tag{4}$$

Since the variance measure from the survey is only proportional to variance, we ignore a and b in this regression and focus on the estimated value of ϕ. Relative to equation (2), ϕ is then an estimate of $\tau\tilde{\rho}$. We emphasize, however, that a nonzero ϕ for the survey in and of itself does not necessarily imply nonrational expectations in the data.[7]

We allow survey expectations to embody signals other than lagged realized variance. In particular, we do not require that the error term is independent and identically distributed (i.i.d.) and compute standard errors using block bootstrap with a six-quarter block length. We set $J = 11$ so that we use one year of lagged realized variance given our finite sample. Our results are not sensitive to small variation in the number of lags. The survey is quarterly, $t = 3, 6, 9, \ldots, T$, but we use the monthly frequency for rv (sum of squared

[7] The lag structure in equation (4) is that of an ARMA(1,1) in realized variance, which is a parsimonious way to capture short- and long-run dynamics in expected variance. Thus, the specification of equation (4) also allows for a commonly used process for the objective dynamics of realized variance. See Section I of the Internet Appendix for further details about the survey, as well as derivations showing how the lag structure in equation (4) relates to an ARMA(1,1) and the above model of beliefs. Internet Appendix may be found in the online version of *The Journal of Finance*.

Table I

Survey Expectations

We fit the actual variance process and survey expectations to an exponential weighted average on past realized variance. That is, we fit $y_{t,t+k} = a + b\Sigma_{i=1}^{J}\phi^{i-1}\sigma_{t-i}^2 + \varepsilon_t$ and report the estimated ϕ, where we set J to 12 periods, and k is the horizon over which investors forecast variance in the survey (one year). We then repeat this procedure replacing σ_t^2 on the left-hand side with the expectation of variance from the survey over the same horizon. A higher ϕ from the expectations data indicates that expectations rely more on variance farther in the past compared to the optimal forecast for volatility. We use the Graham and Harvey CFO survey, which is available quarterly and corresponds to a one year forecast horizon. Standard errors in parentheses are calculated using block bootstrap with a six quarter block length in Panel A and using Newey-West with six quarter lag length in Panels B and C.

Panel A: Dependence on Past Variance (ϕ)

Source	Survey	Future Variance
CFO	0.85***	−0.09
	(0.18)	(0.39)
N	73	73
R^2	0.55	0.21
$t(\phi_{survey} - \phi_{fv})$		(2.04)

Panel B: Surveys, VIX, and Future Variance

	VIX_t^2	Future Variance
CFO_t	1.38***	−0.66
	(0.39)	(1.51)
σ_t^2	0.59***	0.40***
	(0.09)	(0.14)
R^2	0.77	0.18
N	73	73

Panel C: Regression of 10 year Expectations on One-Year Expectations

	Volatility	Variance
1 Year	0.38***	0.25***
	(0.06)	(0.04)
R^2	0.47	0.44
N	68	68
$\tilde{\rho}$	0.98	0.96
CI $\tilde{\rho}$	[0.96,0.98]	[0.92,0.97]

daily returns to the S&P500 in a given month) as our model and later data analysis is at the monthly frequency.

Panel A of Table I reports that the estimated value of ϕ is 0.85 with a standard error of 0.18. Thus, survey respondents effectively take into account many lags of realized variance when forming expectations about future variance. We also report the results from projecting actual realized variance over the next 12 months onto lagged rv using the same functional form on the right-hand side as in equation (4). If the survey expectations are rational, these two projections will yield the same coefficients.[8] In this rational benchmark case, ϕ is estimated to be -0.09. The difference between this estimate and the estimate using the survey expectations is statistically significant at the 5% level as reported in the table. In other words, the full-information rational expectations process for realized variance loads heavily on current variance and less on past variance, while agents' expectations appear sticky in the sense that they assign higher-than-optimal weights to additional lags.

Next, we evaluate the extent to which survey expectations reflect the information contained in the VIX, and vice versa. In Panel B of Table I, we regress the squared VIX on the contemporaneous survey expectation of variance, as well as on the most recent monthly realized variance. Both variables are positive and significant at the 1%-level with an R^2 of 77%. If we instead put next year's realized variance on the left-hand side (rightmost column of Panel B), the survey comes in with an insignificant negative sign, while the current variance comes in positive and strongly significant. Thus, a rational forecast does not load on the survey at all after controlling for the current level of realized variance, while the squared VIX is strongly related to the survey. The survey expectations should show up in the VIX if they capture the expectations of market participants, and the results in Panel B of Table I indicate that they do.

As a more direct way to explore this, we use the VIX^2 in place of the survey as a measure of market expected variance and reestimate our ϕ parameter. We estimate a ϕ of 0.42 (standard error of 0.12) using VIX data from January 1990 to April 2020. Thus, the dependence on past variance that leads agents' expectations to respond too slowly (i.e., underreaction) is present not just in surveys, but also in actual market prices of expected variance.[9] However, the lower value for ϕ suggests that this bias is lower in actual financial market prices.

An additional source of survey evidence on volatility comes from Robert Shiller, who each month since July 2001 asks investors to give their assessment of the probability of a stock market crash over the next six months such as that seen in 1987. While this measure is not as direct as a measure of variance expectations, it does gauge agents' beliefs about risk in the stock market in general. We use this survey as a robustness check and relegate the findings to Internet Appendix Table IA.VI. The Shiller survey produces an estimate of ϕ of around 0.77, which is close to the value found using the CFO survey. The evidence of slow updating toward volatility is consistent with the findings of Boutros et al. (2019) and with the work of Coibion and Gorodnichenko (2015), who document sticky expectations in many other contexts.

[8] To see this, note that $E[E[rv_{t,t+12}|y^t, rv^t]|rv_t, \ldots, rv_{t-11}] = E[rv_{t,t+12}|rv_t, \ldots, rv_{t-11}]$, where y^t is the history of other signals and rv^t is the history of rv.

[9] Other evidence also suggests that agents take action based on their reported beliefs about risk, which is important for survey beliefs to affect prices. Giglio et al. (2019) show that survey data on investor beliefs about risk translate into actions in terms of portfolio allocations, while Ben-David, Graham, and Harvey (2013) find that CFO expectations about volatility translate into investment decisions.

B.2. Persistence of Variance Expectations

Next, we evaluate the persistence of agents' beliefs as implied by the CFO survey. Respondents are asked the same questions about the return distribution at the 10-year horizon. Equation (3) shows how long-run forecasts relate to the short-run forecast under the illustrative model of aggregate expectations given earlier. The 10-year and one-year survey forecasts correspond to *cumulative* 10- and one-year variance expressed in annual terms. This implies that

$$Survey_t^{(10yr)} = \tilde{a} + \tilde{b} \times Survey_t^{(1yr)} + \eta_t,$$

where $\tilde{b} = \frac{\sum_{k=0}^{119} \tilde{\rho}^k}{10 \sum_{k=0}^{11} \tilde{\rho}^k} = \frac{1}{10} \frac{1-\tilde{\rho}^{120}}{1-\tilde{\rho}^{12}}$ and where the error term allows this relation to be inexact.[10]

Panel C of Table I shows that $\tilde{b} = 0.25$ with a standard error of 0.04. Under the belief model above, the implied monthly autocorrelation, $\tilde{\rho}$, is 0.96 with a 95% confidence interval from 0.92 to 0.97. This is much more persistent than the autocorrelation of realized variance, which is only 0.71 in this sample, implying that agents overextrapolate when forecasting future variance. The magnitude of this bias is close to the experimental evidence in Landier, Ma, and Thesmar (2019). These results echo Stein (1989) and Giglio and Kelly (2017), who find overreaction of long-term volatility expectations from options data and variance swap prices, respectively. In particular, they argue that longer term expectations of volatility are too volatile relative to those at short horizons, a form of relative overreaction in long-term expectations.

In the next section, we embed the form of belief bias that we document here into a general equilibrium model and assess the asset pricing implications.

II. The Model

We develop an asset pricing model similar to that in Bansal and Yaron (2004) except we allow the representative investor to have biased beliefs regarding the dynamics of stock return volatility. This simple modification allows the model to account for the empirical evidence discussed above.

Let the objective process for aggregate log dividend growth be given by

$$\Delta d_t = \mu + \sigma_t \varepsilon_t, \tag{5}$$

[10] To see this, note that

$$E_t^S[rv_{t+1} + \cdots + rv_{t+12}] = 12\bar{v} + \sum_{k=0}^{11} \tilde{\rho}^k (E_t^S[rv_{t+1}] - \bar{v}),$$

$$\frac{1}{10} E_t^S[rv_{t+1} + \cdots + rv_{t+120}] = 12\bar{v} + \sum_{k=0}^{119} \tilde{\rho}^k (E_t^S[rv_{t+1}] - \bar{v}).$$

$$\sigma_t^2 = \bar{v} + \rho\left(\sigma_{t-1}^2 - \bar{v}\right) + \omega\eta_t, \tag{6}$$

where σ_t^2 is the realized variance of dividend growth innovations, observed at time t, and ε_t and η_t are uncorrelated i.i.d. standard Normal shocks. Variance is persistent with $0 < \rho < 1$.[11]

We model aggregate beliefs about variance following the evidence on expected variance we establish using survey data in Section II.B. Specifically, the representative agent's expectations of the conditional variance of dividend growth are given by

$$E_{t-1}^S\left[\sigma_t^2\right] = \bar{v} + \lambda x_{t-1}, \tag{7}$$

$$\begin{aligned} x_t &= \phi x_{t-1} + (1-\phi)\left(\sigma_t^2 - \bar{v}\right) \\ &= (1-\phi)\sum\nolimits_{j=0}^{\infty} \phi^j\left(\sigma_{t-j}^2 - \bar{v}\right). \end{aligned} \tag{8}$$

The S superscript on the expectations operator highlights that the expectation is taken under the agent's subjective beliefs. These beliefs are identical to those in equation (2) with the substitutions $\phi = \tau\tilde{\rho}$ and $\lambda = \frac{\tilde{\rho}-\phi}{1-\phi}$, which we make for analytical convenience.[12]If $\phi = 0$ and $\lambda = \rho$, the agent has rational expectations about the volatility dynamics, while if $\phi > 0$, the agent has slow-moving volatility expectations, allowing an exponentially weighted average of past variance to affect the current expectation, as opposed to only the current value as the physical volatility dynamics prescribe. The subjective persistence of variance shocks is increasing in λ, holding ϕ constant. We assume $0 < \phi, \lambda < 1$.

Under agents' beliefs, the shock to variance is

$$\begin{aligned} \omega\eta_t^S &\equiv \sigma_t^2 - \bar{v} - \lambda x_{t-1} \\ &= \rho\left(\sigma_{t-1}^2 - \bar{v}\right) - \lambda x_{t-1} + \omega\eta_t, \end{aligned} \tag{9}$$

where $\rho(\sigma_{t-1}^2 - \bar{v}) - \lambda x_{t-1} = E_{t-1}^P[\sigma_t^2] - E_{t-1}^S[\sigma_t^2] = E_{t-1}^P[\omega\eta_t^S]$ is the mistake made when forecasting variance. A P superscript on the expectations operator means that the expectation is taken under the objective measure. We can thus write the dynamics of x_t under the representative agent's beliefs as

$$x_t = (\phi + (1-\phi)\lambda)x_{t-1} + (1-\phi)\omega\eta_t^S. \tag{10}$$

[11] Equation (6) implies that variance can be negative. For ease of exposition, we follow Bansal and Yaron (2004) and proceed as if σ_t^2 is always nonnegative. In Section II of the Internet Appendix, we show that this simplification is not important for our conclusions by solving a model with Gamma-distributed variance shocks, where variance is guaranteed to always be positive.

[12] To see this note that we can write equation (2) as $E_t^S(\sigma_{t+1}^2) = \bar{v} + (\tilde{\rho} - \phi)\sum_{j=0}^{\infty}\phi^j(\sigma_{t-j}^2 - \bar{v})$ with $\phi = \tau\tilde{\rho}$. Since $\lambda = \frac{\tilde{\rho}-\phi}{1-\phi}$, then $\tilde{\rho} - \phi = \lambda(1-\phi)$, and making this substitution gives equations (7) and (8).

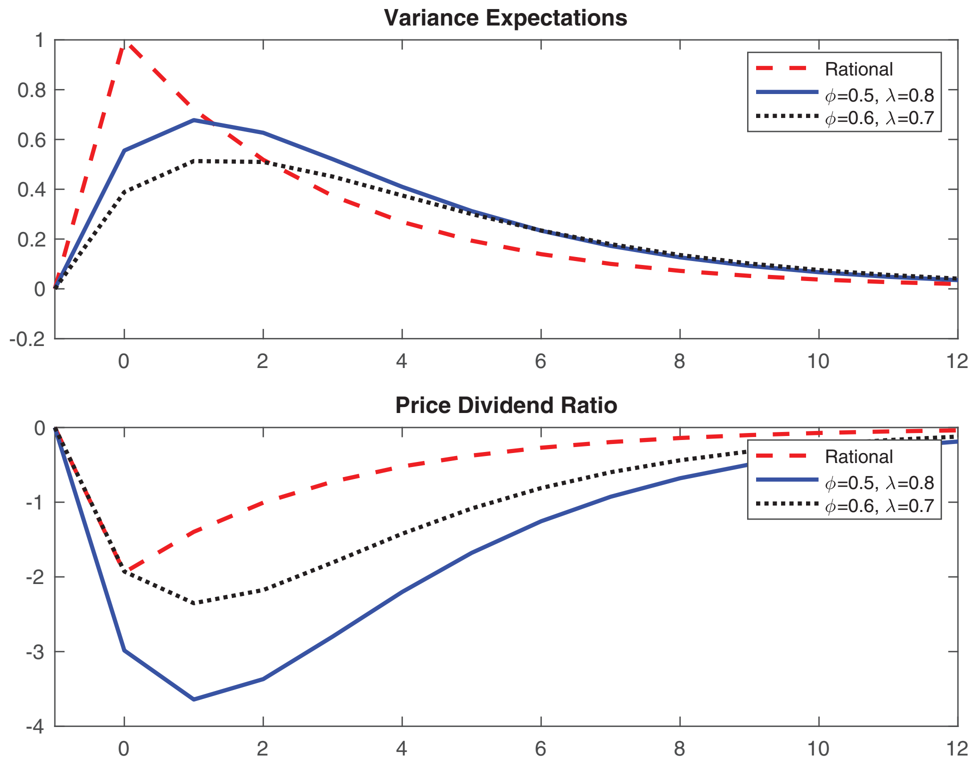

Figure 1. Dynamics of variance expectations in the model. We plot the behavior of agents' expectations of volatility in our model (blue line) and the true path of expected volatility (dashed red line) in response to a one-standard-deviation increase in variance in our model. The dot-dashed black line provides an alternative calibration when we set the scale of expectations lower. Because agents effectively take a weighted average of past volatility, they initially underreact and then subsequently overreact. The variance risk premium then reflects the difference between agents' expectations of volatility minus the rational forecast of volatility, and thus, it goes negative initially and then becomes positive. The bottom panel shows the behavior of stock prices (the price-dividend ratio) that also responds slowly in the model, reflecting agents' slow-moving beliefs. This slow response makes it appear as if equity risk premiums do not initially rise (and potentially even fall) after an increase in volatility but then rise later after the volatility shock has largely subsided. The x-axis is in months. (Color figure can be viewed at wileyonlinelibrary.com)

Note that variance expectations are sticky relative to the true variance dynamics if $\phi > 0$ and $\lambda \geq \rho$, as the persistence of x_t is then higher than the true persistence of σ_t^2 (that is, $\phi + (1-\phi)\lambda > \rho$). Also note that the shock itself is moderated by a factor of $1-\phi$. Figure 1 shows the impulse response from a positive variance shock (η_0) for objective and subjective expected variance. The parameter values are calibrated to the data as described below. The true AR(1) dynamics of variance are reflected in the monotonically decaying response in the rational case (dashed red line). The solid blue line gives the impulse response of agents' expected variance as reflected in the dynamics of x_t. Agents initially underreact, as ϕ is greater than zero in this case, but the higher persistence of x_t leads to subsequent overreaction.

We assume a representative stockholder with consumption equal to aggregate dividends whose marginal utility prices all claims in the economy. Following Bollerslev, Tauchen, and Zhou (2009), the agent has Epstein-Zin utility (Epstein and Zin (1989)), where β, γ, and ψ are the time-discounting, risk aversion, and intertemporal substitution parameters, respectively. The stochastic discount factor is therefore

$$M_t = \beta^\theta e^{-\frac{\theta}{\psi}\Delta d_t + (\theta-1)r_t}, \tag{11}$$

where $\theta = \frac{1-\gamma}{1-1/\psi}$ and r_t is the log return to the aggregate dividend claim. We use the standard log-linearization techniques of Campbell and Shiller (1988) and Bansal and Yaron (2004) to derive equilibrium asset prices (see Section I of the Internet Appendix for details). In particular, we assume that aggregate log returns are $r_t = \kappa_0 + \kappa pd_t - pd_{t-1} + \Delta d_t$, where pd is the aggregate log price-dividend ratio and κ is a constant close to but less than one that arises from the log-linearization. We then obtain

$$pd_t = c - Ax_t, \tag{12}$$

where $A = -\frac{1}{2}\frac{\lambda(1-\gamma)(1-1/\psi)}{1-\kappa(\phi+(1-\phi)\lambda)}$. Notice that if $\gamma, \psi > 1$, we have that $A > 0$. This is the standard preference parameter configuration for asset pricing models with Epstein-Zin preferences. It implies that the price-dividend ratio is low when agents perceive variance to be high, as in the data.

A. Equity Risk Premium Dynamics

Let r_t and $r_{f,t}$ denote the aggregate log return and risk-free rate in period t, respectively. The subjective conditional risk premium of log returns in this economy is

$$E^S_{t-1}[r_t - r_{f,t}] = \left(\gamma - \frac{1}{2}\right) E^S_{t-1}[\sigma_t^2] + \delta_r, \tag{13}$$

where δ_r is a constant given in the Internet Appendix (Section I) that captures the price effect of discount rate shocks due to the variance shocks (η_t). The first term reflects the standard risk-return trade-off that is linear in the conditional variance of dividend growth, where the $-1/2$ part arises as this is the log return risk premium.

The conditional variance of log returns is determined by both the conditional variance of dividend growth and the impact of the variance shock on the price-dividend ratio

$$\mathrm{Var}^S_{t-1}(r_t) = \Theta + E^S_{t-1}[\sigma_t^2], \tag{14}$$

where $\Theta = (\kappa A(1-\phi)\omega)^2$.

The *objective* risk premium, however, is

$$E_{t-1}^{P}\left[r_t - r_{f,t}\right] = E_{t-1}^{S}\left[r_t - r_{f,t}\right] + \kappa(1-\phi)A\Big(E_{t-1}^{S}\left[\sigma_t^2\right] - E_{t-1}^{P}\left[\sigma_t^2\right]\Big), \tag{15}$$

where the P superscript on the expectation denotes that it is taken using the true, objective variance dynamics. Note that the risk premium loads *negatively* on true conditional variance, as $\kappa(1-\phi)A > 0$ for our calibrated parameters—a major departure from earlier literature. To see where equation (15) comes from, recall that the shock to agents' beliefs about variance is predictable (see equation (9)). The mistake is persistent, which magnifies its effect on prices as given by the term $\kappa A(1-\phi)$.[13]

The mistakes in conditional variance expectations are reflected in current discount rates and therefore prices. Consider a positive shock to variance ($\eta_{t-1} > 0$). With $\phi > 0$, expectations are sticky, which means that investors do not update their beliefs sufficiently and initially underreact to the variance shock. Thus, $E_{t-1}^{P}[\sigma_t^2] > E_{t-1}^{S}[\sigma_t^2]$. Since $A > 0$ in the relevant calibrations, this means that if the mistake is sufficiently large, a *positive* shock to variance can *decrease* next period's objective risk premium. The reason is that investors will on average perceive a positive shock to discount rates next period as the realized value of σ_t^2 on average being higher than they had expected. This leads to a predictable decline in the price-dividend ratio under the objective measure. Recall that the price-dividend ratio is given by $c - Ax_t$. Hence, the price-dividend ratio falls at the impulse, but note that it keeps falling in the following period due to the increase in discount rates when agents learn that variance is higher than expected and $E_{t-1}^{S}[\sigma_t^2]$ rises. Subsequently, given the too-persistent variance expectations, investors eventually overreact to the volatility shock, which leads to $E_{t+j-1}^{P}[\sigma_{t+j}^2] < E_{t+j-1}^{S}[\sigma_{t+j}^2]$ for some $j > 0$. In this case, the second term in equation (15) becomes positive and the conditional risk premium overshoots.

Upon impact, a positive shock to variance decreases prices as the long-run impact on discount rates is positive when A is positive. This is consistent with the negative contemporaneous correlation of realized variance and returns in the data (e.g., French, Schwert, and Stambaugh (1987)). In particular, shocks to returns are

$$r_t - E_{t-1}^{P}[r_t] = -\Theta^{1/2}\eta_t + \sigma_t\varepsilon_t, \tag{16}$$

where $\Theta^{1/2} = \kappa A(1-\phi)\omega$ encodes the present value impact of the shock to variance (η_t) due to its effect on the discount rates that agents require for holding the risky asset.

[13] This expression is found by using the Campbell-Shiller return approximation and noting that $E_{t-1}^{P}(-\kappa pd_t) - E_{t-1}^{S}(-\kappa pd_t) = -\kappa A(E_{t-1}^{P}[x_t] - E_{t-1}^{S}[x_t]) = -\kappa A(1-\phi)(E_{t-1}^{P}[\sigma_t^2] - E_{t-1}^{S}[\sigma_t^2])$.

B. Variance Risk Premium Dynamics

In addition to the equity claim, we also price a variance claim with payoff

$$RV_t \equiv \Theta + \sigma_t^2, \tag{17}$$

where RV_t stands for realized variance at time t.[14] We define the time $t-1$ implied variance (IV_{t-1}) as the swap rate that gives a one-period variance swap a present value of zero

$$0 = E_{t-1}^S[M_t(RV_t - IV_{t-1})]. \tag{18}$$

Thus

$$IV_{t-1} = E_{t-1}^S\big[R_{f,t}M_tRV_t\big]. \tag{19}$$

As is standard in the literature, we define the (objective) expected payoff of a position in the variance swap where one is paying the realized variance and receiving the implied variance as the variance risk premium

$$VRP_{t-1} = IV_{t-1} - E_{t-1}^P[RV_t]. \tag{20}$$

The dynamics of this risk premium depends directly on variance expectations per equation (19).

The equilibrium implied variance is

$$IV_{t-1} = E_{t-1}^S[RV_t] + \delta_{IV}, \tag{21}$$

where $E_{t-1}^S[RV_t] = E_{t-1}^S[\Theta + \sigma_t^2] = \Theta + \bar{v} + \lambda x_{t-1}$ and $\delta_{IV} = (\frac{1}{2}\gamma^2 - \frac{1/\psi-\gamma}{1-1/\psi}\kappa(1-\phi)A)\omega^2$. The second term is an unconditional risk premium required by agents

[14] Our model definition of realized variance is motivated by industry practice for variance swap payoffs, where monthly realized variance is the sum of squared daily log returns within the month. In the model, squared monthly log returns are given by

$$\big(r_t - E_{t-1}[r_t]\big)^2 = \Theta\eta_t^2 + 2\sigma_t\Theta^{0.5}\eta_t\varepsilon_t + \sigma_t^2\varepsilon_t^2.$$

To approximate the use of higher frequency data to estimate realized variance within our model, we assume that the second moments of realized shocks equal their continuous-time limit. Setting $\eta_t^2 = \varepsilon_t^2 = 1$ and $\eta_t\varepsilon_t = 0$ in the above gives the realized variance in equation (17). In benchmark equilibrium models, typically calibrated at the monthly frequency (e.g., Bollerslev, Tauchen, and Zhou (2009), Drechsler and Yaron (2011)), there is no clear counterpart to this multifrequency approach where IV and RV are monthly but where RV is estimated using daily data. In the models cited above, the definition of the IV_t is the risk-neutral expectation of the market return variance in month $t+2$. For example, IV at the end of January is the risk-neutral expectation at the end of January of market return variance in March. We define RV in a manner that avoids this one-month offset that is at odds with the data definitions. This brings the model closer to the moments from the data that we use for calibration of the model parameters. While it is convenient to align the model definitions more closely to the timings used in the data, we note that our model results would also go through with alternate definitions of the variance risk premium used in the earlier literature.

due to the variance claim's exposure to variance shocks. The conditional variance risk premium is then

$$\begin{aligned} VRP_{t-1} &= IV_{t-1} - E^P_{t-1}[RV_t] \\ &= \delta_{IV} + E^S_{t-1}[RV_t] - E^P_{t-1}[RV_t] \\ &= \delta_{IV} + E^S_{t-1}\left[\sigma_t^2\right] - E^P_{t-1}\left[\sigma_t^2\right]. \end{aligned} \tag{22}$$

Thus, the dynamics of the variance risk premium share a component of the dynamics of the equity risk premium (equation (15)), namely, the mistakes that agents make in their variance expectation. Thus, agents will initially underreact to the variance shock but subsequently overreact due to their sticky expectations, which leads to time variation in the variance risk premium similar to that in the data. In fact, the lagged variance risk premium forecasts equity returns, as it does in the data and as it does in Bollerslev, Tauchen, and Zhou (2009). However, in their model, this is due to time-varying variance of variance, which we abstract from in this baseline version of our model.

We next calibrate the parameters of the model to assess if it can quantitatively account for the empirical observations discussed earlier.

C. Model Calibration

We calibrate the model to moments that are at the heart of the issues we seek to address with the model. The data are monthly and run from 1990 through April 2020. We use the VIX_t^2 to proxy for IV_t, where VIX_t is the option-implied risk-neutral volatility of stock returns over the next month. We calculate RV_t as the sum of daily squared log excess market returns in month t.

Panel A of Table II gives the parameters of the baseline model. We match the mean, autocorrelation, and variance of RV_t in the data with the parameters governing the objective variance dynamics in the model ($\bar{v}$, ρ, and ω). We set the risk-aversion parameter γ by matching the equity premium and we take the elasticity of substitution ψ to be 2.2 as estimated in Bansal, Kiku, and Yaron (2016). Finally, we set ϕ to match the response of the variance risk premium (VRP) to a shock to RV in the model to that in the data, and we set λ to match the variance of IV_t in the model to the variance of the VIX_t^2 in the data. We set $\kappa = 0.97^{1/12}$, consistent with values used in earlier literature and the average level of the price-dividend ratio in the data. Our moments of interest do not require that we estimate the time-discounting parameter, β, or the mean of dividend growth, μ.[15] Table II gives the parameter values as well as the moments from the data used in the calibration. Note that the chosen value of ϕ is conservative in the sense that it is lower than that estimated using the survey data. We choose a ϕ of 0.5 that is between the estimate from the survey

[15] This is why we set the log-linearization parameter κ exogenously to a standard value in the literature. In our monthly calibration, κ is very close to one and there is little sensitivity to reasonable variation in this parameter to the moments we target.

Table II

Calibration

Panel A reports the calibrated values for the parameters, along with a description of the corresponding moment used in the calibration. Panel B gives the values of selected moments in the model and in the data, along with a description of the moment. Here, RV_t is the realized variance in month t, VIX_t is the end of month t value of the VIX (not annualized), and $\alpha((c/RV_{t-1})r_{m,t}, r_{m,t})$ is the annualized intercept of a monthly regression of a volatility-timed version of the market portfolio on the market. The timing strategy has time $t-1$ weight c/RV_{t-1}, where c is a constant that makes the volatility of this strategy the same as that for a static position in the market. Finally, $\tilde{RV}_t$ denotes the shock to realized variance obtained from an AR(1) process.

Panel A: Parameters

Parameter	Description	Value	Targeted Moment(s)
γ	Risk Aversion	3	Equity Premium
ψ	Elasticity of Intertemporal Substitution	2.2	Literature/VRP
$\bar{v}$	Unconditional Variance (Monthly)	0.25%	Data
ρ	Persistence of Variance	0.71	Data
ω	Volatility of Variance Shocks (Monthly)	0.31%	Data
ϕ	Expectation Stickiness	0.5	VIX/Surveys
λ	Scale of Expectations	0.8	Vol of VIX/Surveys

Panel B: Moments

Moment	Description	Model	Data
$E[r_{m,t} - r_{f,t}]$	Equity Premium (Annual)	7.9%	7.7%
$\sqrt{E[RV_t]}$	Square Root Avg. Variance (Annual)	18%	18%
$\rho(RV_t, RV_{t-1})$	Persistence of Variance (Monthly)	0.71	0.71
$\sigma(RV_t)$	Volatility of Variance (Monthly)	0.44%	0.44%
$\sigma(VIX_t^2)$	Volatility of VIX^2 (Monthly)	0.33%	0.35%
$\rho(VIX_t^2, RV_t)$	Correlation RV and VIX^2 (Monthly)	0.88	0.85
$\rho(VIX_t^2, VIX_{t-1}^2)$	Persistence of VIX^2 (Monthly)	0.91	0.84
$\alpha(\frac{c}{RV_{t-1}} r_{m,t}, r_{m,t})$	Volatility-Managed Alpha (Moreira Muir)	1.4%	4.9%
$\rho(r_{m,t}, \widetilde{RV}_t)$	Correlation of Returns and Vol Shocks	−0.20	−0.38

data (0.87) and the estimate we get if we use the VIX (0.4) in place of surveys as a measure of market variance. Our choice of λ implies a persistence of volatility dynamics under agents' beliefs of $\phi + (1-\phi)\lambda = 0.9$, which is again close to but more conservative than the survey estimates based on long-term and shorter-term volatility expectations (0.96).

Figure 1 shows the impulse response of a one-standard-deviation shock to RV under rational beliefs (dashed red line) and the calibrated subjective beliefs about next period's variance (solid blue line). We show an alternative calibration in the black line that gives intuition for a lower value of ϕ. The subjective beliefs display a hump-shaped "slow-moving" response due to the beliefs loading on too many lags of variance and not enough on current variance. The difference between the rational and the subjective beliefs—initial underreaction followed by longer term overreaction—is what gives rise to deviations from the standard models in terms of the observed risk-return trade-off. The figure also shows how λ scales the impulse response (see dotted black line versus solid blue line). The bottom panel plots the price-dividend ratio. Notably, in our model, calibration prices initially fall by more than in the rational case—this is because of the overextrapolation leading expected variance to be too persistent. However, consistent with underreaction, prices can continue to fall after the initial shock, breaking the standard risk-return trade-off. In this way, we match the stylized fact that realized returns fall substantially when variance increases (consistent with a discount rate effect) yet next period's returns are not high on average.

III. Model Comparison to Stylized Facts

A. Stylized Facts

We outline the main stylized empirical facts from the literature that we target in our model, which we extend using more recent data. Since most of these facts are already documented in the empirical literature, we relegate discussion of robustness to Section V of the Internet Appendix.

Data and Sources. We study U.S. data from December 1991 to April 2020 for which we have stock market excess returns, the VIX (taken as the VIX on the last day of the month, thus representing forward-looking volatility for the month), realized variance (computed as the sum of squared daily log returns within a month), and a measure of expected variance. Stock return data use the return on the S&P500 index over the risk-free rate taken from Ken French. In addition, we study variance swap returns, VIX futures returns, and straddle returns to capture claims on future volatility or variance from several sources. Our main data source for variance swap returns is Dew-Becker et al. (2017), who provide variance swap returns based on dealer quotes from 1996 to 2017. VIX futures returns come from Cheng (2018) for 2004 to 2017 and we supplement these data with returns on the VXX exchange-traded fund (*ETF*) from 2017 to 2020 (as Cheng (2018) notes the VXX ETF tracks VIX futures returns and has a correlation near one with one-month VIX futures in the overlapping sample). We further measure the variance risk premium (*VRP*) as the squared VIX minus realized variance as in Bollerslev, Tauchen, and Zhou (2009). When using returns (e.g., variance swap or VIX futures), we take the negative of the returns, so our measure is the return for selling variance or being short the VIX. This means that the unconditional premium for these returns is positive as the exposure to volatility is negative. Finally, we obtain data from Johnson (2017), who provides daily straddle returns and synthetic variance swap returns from underlying options that we cumulate to monthly returns from 1996 to 2019. These data have the advantage of avoiding over-the-counter prices or quotes that may have illiquidity concerns.

Expected Variance. To get a measure of objective expected variance at time $t-1$, which we denote by $\hat{\sigma}^2_{t-1}$, we use high-frequency intraday data on the S&P500 and follow the HAR model used in Bekaert and Hoerova (2014). This model uses squared five-minute returns at the daily, weekly, and monthly horizons to forecast next month's realized variance. Importantly, our forecast is conducted in real time so that at time $t-1$ the forecast only uses information up to $t-1$. Our intraday data are available in 1990 but we use a two-year burn-in period to construct the expected variance forecast. Further details are contained in Section IV of the Internet Appendix. Following our model, we proxy for subjective expected variance as a weighted average of past expected variance over the past six months, denoted by $\frac{1}{\Sigma_{j=1}^{6}\phi^j}\Sigma_{j=1}^{6}\phi^j\hat{\sigma}_{t-j}^2$. The term in front of the sum ensures that weights sum to one. We use $\phi = 0.5$ as in the model calibration, and only six lags since weights ϕ^j are close to negligible beyond this.

Empirical Results. Table III, Panel A, shows forecasting regressions for volatility claims on current expected variance and the exponential weighted average of past expected variance. In the first five columns, for each of the (short) volatility dependent returns, we see a negative coefficient on expected variance and a positive coefficient on the weighted average term, with the magnitude of the coefficients roughly similar. That is, in all cases, the regression emphasizes the difference between the weighted average and objective expected variance as the model predicts. Thus, when agents' beliefs about variance are high relative to an objective forecast of variance, volatility claims are "expensive" in the sense that being short volatility or variance is profitable over the next month. The monthly R^2 implies a reasonably high degree of predictability for all volatility claims and the coefficients are significant in each of the columns.[16] These results are closely related to those in Cheng (2018), who shows that increases in expected volatility negatively predict the VIX futures return. Following Cheng (2018), in the first column for VIX futures returns, we use expected volatility in the place of expected variance, and the weighted average of expected volatility. Using variance leads to similar results but allows for less of a direct comparison to Cheng (2018). While the magnitudes of coefficients in columns (5) to (7) are easily interpretable in our model, those in columns (1) to (4) are less so. For straddle returns, the coefficient on the difference between expected variance and a weighted average of expected variance (the "mistake" in beliefs) is about −30. The standard deviation of the difference is 0.16% per month, so a one-standard-deviation increase in the difference is equivalent to nearly a 5% change in the straddle risk premium. This is large given the unconditional monthly average return of 9.5% for straddles. Similarly, for variance swaps, a one-standard-deviation increase in the difference is equivalent to about a 9% change versus the unconditional premium of 25%.

Column (6) shows that only expected variance predicts future realized variance, while the additional lags of expected variance do not. The coefficient is not statistically different from one and the R^2 is high at 47%, confirming that our real-time expected variance forecast predicts future variance well (when

[16] One concern with these predictive return regressions is the bias documented by Stambaugh (1999). However, this issue is negligible in our setting because the volatility mistake (difference between weighted average and objective expected variance) that forecasts returns has a low monthly autocorrelation of around 0.28.

Table III
Stylized Facts

Panel A runs predictive regressions of variance dependent returns, the VIX, and future realized variance on expected variance ($\hat{\sigma}^2_{t-1}$) and a weighted average of expected variance over the past six months $(1-\phi)\Sigma_{k=1}^{6}\phi^{k-1}\hat{\sigma}^2_{t-k}$, where $\hat{\sigma}^2_{t-1}$ represents expected variance at time $t-1$, while σ^2_t is the realized variance of daily market returns in month t. Variance-dependent returns are short positions in VIX futures, straddles, and variance swaps. Theses returns represent the premium for insuring against future increases in VIX, variance, or volatility (so that the variance risk premium is positive on average). Panel B runs excess stock returns (market returns over the risk-free rate) on expected variance, the average of past expected variance, and the implied variance from the VIX. Data are monthly from 1992 to 2020. The variance swap, VIX futures, and straddle return data are from 1996 to 2019, 2004 to 2020, and 1996 to 2020, respectively. Standard errors in parentheses use Newey-West correction with 12 lags.

Panel A: Variance Returns

	Vix Fut (1)	Straddle (2)	Var Swap (3)	Var Swap (TJ) (4)	$VIX^2_{t-1}-\sigma^2_t$ (5)	σ^2_t (6)	VIX^2_{t-1} (7)
$\hat{\sigma}^2_{t-1}$	−4.99*****	−29.18**	−40.62***	−55.90***	−1.26***	1.53***	0.30**
	(1.73)	(11.49)	(14.16)	(14.52)	(0.49)	(0.54)	(0.12)
$\Sigma_{j=1}^{6}\phi^j\hat{\sigma}^2_{t-j}$	5.05**	36.57**	42.90**	58.96***	1.23***	−0.58	0.64***
	(2.07)	(14.24)	(20.19)	(16.86)	(0.46)	(0.46)	(0.07)
N	194	292	264	282	334	334	335
R^2	0.0521	0.0240	0.00682	0.00496	0.172	0.469	0.790

Panel B: Stock Market Returns

	Market (1)	Market (2)	Market (3)	Market (4)
$\hat{\sigma}^2_{t-1}$	−0.59		−2.21	−3.44***
	(0.80)		(1.38)	(0.87)
VIX^2_{t-1}		0.41		3.69***
		(1.16)		(1.25)
$\Sigma_{j=1}^{6}\phi^j\hat{\sigma}^2_{t-j}$			2.17	
			(2.11)	
N	340	363	335	340
R^2	2.04e−05	−0.00170	0.00105	0.0211

we run this as a univariate regression, without the weighted average, we obtain a coefficient of 1.10 and standard error of 0.20). Column (7) shows that market-implied variance (the squared VIX), however, loads substantially on the weighted average of past expected variance in addition to current expected variance as our model predicts. The differential pattern in columns (6) and (7) helps explain the patterns in the first five columns. The R^2 on the VIX is very high, indicating that we capture the majority of VIX variation from these two components. While we interpret all of these results through biases in beliefs as in our model, an alternative interpretation is through rational risk premiums. However, as noted in Cheng (2018), this interpretation is not natural because volatility claims are typically much riskier when expected volatility is high relative to the past. In Internet Appendix Table IA.V, we run these return regressions at a weekly frequency to assess where the monthly return predictability is most concentrated. Consistent with Cheng (2018), the predictability is stronger in the first two weeks relative to weeks 3 and 4.[17] This is consistent with our model because underreaction in our model is most pronounced in the near term.

Panel B shows the risk-return trade-off regressions of the market excess return on expected variance and the VIX. Column (1) shows that expected variance has no predictive power for future returns, so the risk return trade-off is weak, echoing a much longer literature on this finding (Glosten, Jagannathan, and Runkle (1993), Whitelaw (1994), Lettau and Ludvigson (2010)).[18] We note that these papers come to this conclusion over a variety of sample periods. This result also holds when using the VIX to forecast market returns in column (2). However, as column (4) shows, the variance risk premium ($VIX_{t-1}^2 - \hat{\sigma}_{t-1}^2$) predicts returns with a positive sign, while neither the VIX or the RV predict returns by themselves. This confirms results from Bollerslev, Tauchen, and Zhou (2009) in our extended sample. However, we note that the first two columns are puzzling from the perspective of the model in Bollerslev, Tauchen, and Zhou (2009), in which the VIX alone is a strong predictor of returns, because it embeds the variance risk premium and because it reflects expected future variance, both of which strongly contribute to the equity risk premium in that model.

These facts are in line with findings from the empirical literature and this table should be mainly viewed as extending them in a more recent sample. As such, we leave extensive robustness checks to Section V of the Internet Appendix. We run subsample analysis focusing on the financial crisis, we use volatility in place of variance to reduce dependence on high-variance observations, and we run weighted least squares to downweight high-volatility periods that might overly influence our results. An important takeaway is that high-variance periods are important, especially for the results in Panel B predicting equity market returns, while results predicting the variance claim returns (first four columns of Panel A) appear more robust. The latter evidence is true in our model as well—variance return predictability identifies our main mechanism more sharply while the equity return is exposed to additional sources of shocks. These results are echoed in Johnson (2019), who argues that the evidence for the variance risk premium predicting stock returns is weaker than previously recognized. However, Bollerslev et al. (2014) study the ability of the variance risk premium to predict returns across eight different countries and argue that it is a robust feature of the data. We provide international evidence along these lines in Internet Appendix Table IA.VIII. In a review article,

[17] See (Cheng, 2018, figure 4).

[18] See, for example, Brandt and Kang (2004), Moreira and Muir (2017), Moreira and Muir (2019), and Eraker (2020).

Zhou (2018) finds that the variance risk premium helps predict returns across many asset classes including stocks, credit, currencies, and bonds and provides many additional references, suggesting that the variance risk premium helps predict equity returns (Drechsler and Yaron (2011), Bollerslev, Todorov, and Xu (2015)).

B. Stylized Facts in the Model

We report stylized facts in our calibrated model in Table IV. To emphasize intuition and the importance of the bias to match the data, we report results as we vary the parameter ϕ. We consider the fully rational case in the model as a benchmark ($\phi = 0, \lambda = 0.71$) in column (2), and we use our calibration of $\lambda = 0.8$ in the remaining columns while increasing ϕ to 0.3, 0.5, and 0.8.

We provide results on the relation between risk and returns (regression of future market returns on current and past variance), volatility-managed alphas, the correlation between realized returns and variance shocks, the forecasting regressions of stock returns using the variance risk premium, the relation between the conditional variance risk premium and current and past variance, and the correlation of the model-implied variance (VIX^2) and realized variance.

We first note that the dependence of future returns on current variance declines as we increase ϕ, while the dependence on past variance (ϕ weighted average) increases as we increase ϕ. The rational case in our model implies that only current variance should predict returns, with zero weight on the past average, consistent with the basic risk-return trade-off intuition. With high enough ϕ, current variance can have zero or even negative relation to next-period returns, while the average of past variance comes in positively for larger values of ϕ. These results are mirrored in the next row that documents the volatility-managed alpha, with empirical numbers taken from Moreira and Muir (2017). The alpha is positive in the data, reflecting a weak risk-return trade-off. As we increase ϕ and the risk-return trade-off weakens, we increase the volatility timing alpha as well.

The contemporaneous correlation between realized returns and shocks to variance does not depend too strongly on ϕ, and quantitatively decreases slightly as we increase ϕ. The reason is that there are two effects that run in opposite directions in our model: a higher ϕ implies a lower reaction to volatility news through slow-moving expectations, but also leads agents' expectations to be more persistent than the true volatility process. This second effect results in a larger discount rate response to volatility shocks as they last longer in agents' expectations, and tends to move prices more when volatility changes, while the first effect dampens the response to volatility news. Thus, our model retains the negative correlation between returns and variance shocks even when underreaction to volatility is large.

We next turn to implications for the variance risk premium. First, the variance risk premium forecasts stock returns strongly in the data, which the

Table IV

Stylized Facts in Model and Sensitivity to ϕ

We compare our main facts from Table III in the data (first column) versus our model (remaining columns). The risk-return trade-off regresses one-month-ahead market excess returns on current variance and an average of variance over the past six months. The volatility-managed alpha is taken from Moreira and Muir (2017) based on their volatility timing strategy (see text for details). All other data are monthly from 1990 to 2020 in Table III. We show how our results change as we increase the parameter ϕ across columns. The second column is the rational model case, $\phi = 0$, $\lambda = 0.71$, while in the other columns, we use our calibrated value of $\lambda = 0.8$.

	Data	Model			
	(1)	(2)	(3)	(4)	(5)
		Rational Case	$\phi = 0.3$	$\phi = 0.5$	$\phi = 0.8$
		Risk-Return Trade-Off			
σ^2_{t-1}	−2.21	2.71	−0.23	−1.52	−1.05
	(1.38)				
$\Sigma_{j=1}^{6}\phi^j\hat{\sigma}^2_{t-j}$	2.17	0	3.02	3.74	1.78
	(1.42)				
R^2	0.1%	5.3%	3.5%	3.4%	1.4%
		Volatility-Managed Alpha			
α	4.86	−0.09	0.30	1.40	5.33
	(1.56)				
		Correlation: Realized Returns and Vol Shocks			
	−0.38	−0.24	−0.20	−0.16	−0.15
		Forecasting Returns with Variance Risk Premium			
VRP_{t-1}	3.47	0	7.40	2.58	1.42
	(0.92)				
R^2	2.3%	0%	1.9%	1%	0.5%
		Expected Variance Risk Premium ($VRP_{t-1} = VIX^2_{t-1} - E_{t-1}[\sigma^2_t]$)			
σ^2_{t-1}	−1.26	0	−0.25	−0.70	−0.92
	(0.49)				
$\Sigma_{j=1}^{6}\phi^j\hat{\sigma}^2_{t-j}$	1.23	0	0.39	0.80	0.58
	(0.46)				
		Correlation: VIX^2 and Realized Variance			
	0.86	1	0.98	0.92	0.73

model can account for when $\phi > 0$. The variance risk premium itself (here measured as VIX^2 minus a forecast of realized variance based on current and past variance) is negatively related to current variance and positively related to past variance. This is just the result from Table I that, relative to future variance, VIX loads more on past variance and less on current variance. The model can generate this pattern with $\phi > 0$ (which is required for the variance risk premium to have time variation). As ϕ increases, so that expectations are slow moving, current variance forecasts this premium more negatively and past variance forecasts the premium more positively. This is simply because the mistake in expectations is larger as we increase ϕ.

Finally, empirically, there is a strong correlation between implied variance (VIX^2) and realized variance (0.86). If VIX is influenced by beliefs about variance, this suggests that such beliefs are highly correlated with an objective forecast. In the model, this correlation weakens as we increase ϕ, as it implies investors make larger mistakes. Notably, this correlation remains fairly high even for large values of ϕ. This may seem surprising, since it implies that the subjective forecast of variance is strongly correlated with the objective measure in the model, which means that mistakes are actually fairly small, even when we increase ϕ. But note that volatility is persistent, and agents' beliefs still put most weight on recent variance. Because volatility is fairly persistent, putting weight on lagged variance results in only a modest mistake, and these weights decay fairly quickly for longer lags (which have weight ϕ^k). This is an important point since it highlights that while the degree of bias in our model may appear large, persistence in variance actually implies only modest mistakes and, as we show shortly, modest profits from trading. Only in the case in which ϕ is highest at 0.8 is this correlation in the model lower than what we see empirically.

Having discussed this intuition, we note that ϕ equal to approximately 0.5 does a fairly good job jointly accounting for the facts in the data in terms of the risk-return trade-off, volatility-managed alpha, variance risk premium dependence on past variance, and correlation between VIX and realized variance. However, these results also suggest some tension in the model in terms of jointly matching all facts quantitatively. In particular, the risk-return trade-off is even weaker in the data than in the model with $\phi = 0.5$, and hence, a large ϕ is needed to match this moment. On the other hand, the variance risk premium results favor a more modest value of ϕ for the magnitudes of the variance risk premium on past variance to not be too large. Most importantly, however, the model with biased beliefs matches the moments better on balance than the rational benchmark.

C. Volatility-Managed Portfolios

Moreira and Muir (2017) document that volatility-managed factor portfolios yield positive alpha in standard (Gibbons, Ross, and Shanken (1989)) type return regressions. For the market factor, they consider a strategy that each period has a portfolio weight in the market that is inversely proportional to RV. They show that the alpha of such a strategy relative to the buy-and-hold market factor can be approximated by

$$\alpha = -\frac{c}{E[RV_t]}\mathrm{Cov}\left(E_t[RV_{t+1}], \frac{\mu_t}{E_t[RV_{t+1}]}\right), \tag{23}$$

where $E_t[r_{t+1} - r_{f,t}] = \mu_t$ and where c is a constant that scales the timing portfolio to have the same return variance as the market. Since there is no strong risk-return trade-off in the model with biased beliefs, the covariance above is negative, which gives rise to a positive alpha as in the data. Our simple variance process allows negative values for variance, and therefore to calculate this covariance we use the approximation

$$\frac{\mu_t}{E_t[rv_{t+1}]} \approx \frac{\bar{\mu}}{\bar{v} + \Theta} + \frac{1}{\bar{v} + \Theta}(\mu_t - \bar{\mu}) - \frac{\bar{\mu}}{(\bar{v} + \Theta)^2}(E_t[rv_{t+1}] - \bar{v} - \Theta), \tag{24}$$

and report the alpha for the volatility-managed market portfolio in Table II as[19]

$$\alpha \approx -0.6 \times \frac{1}{\bar{v} + \Theta} Cov\left(E_t[rv_{t+1}], \mu_t - \frac{\bar{\mu}}{\bar{v} + \Theta} E_t[rv_{t+1}]\right), \tag{25}$$

which is equal to 1.4% annualized—on the same order of magnitude as the 4.9% that Moreira and Muir (2017) document. As we emphasize in the prior section, to fully match this alpha would require a larger value of ϕ.

D. Comparison to Bollerslev, Tauchen, and Zhou (2009)

Standard asset pricing models typically struggle with the facts outlined above because they suggest that an increase in risk (volatility) will be associated with heightened risk premiums at all horizons. This relationship will be strongest in the near term but decays with horizon as volatility is mean-reverting. For example, Moreira and Muir (2017) show that the risk-return trade-off in leading models is strong, including models with habit formation (Campbell and Cochrane (1999)), long-run risk (Bansal and Yaron (2004), Drechsler and Yaron (2011)), rare disasters (Barro (2006), Wachter (2013)), and intermediary models (He and Krishnamurthy (2013)). Further, expected returns will typically rise most on impact and will gradually fade over time as volatility fades.

Bollerslev, Tauchen, and Zhou (2009; BTZ hereafter) provide a rational benchmark model of the dynamics of the variance risk premium and the conditional equity premium and therefore overlap with some of the stylized facts we seek to explain. In this model, the representative agent has rational expectations and the volatility of volatility follows a mean-reverting process. The time variation in the amount of variance risk gives rise to time-varying expected returns to variance swaps and the market risk premium. In particular, the lagged variance risk premium in their model predicts future excess market returns as in the data. To highlight how the subjective beliefs model of this paper

[19] Moreira and Muir (2017) find that $\frac{c}{E[RV_t]} \approx 0.6$ in the data, and we use this value to compute the volatility-timed portfolio alpha implied by our model.

Table V

BTZ Comparison

We compare our model to the model of Bollerslev, Tauchen, and Zhou (2009), given in column (2). The text discusses the calibration of BTZ. The remaining columns show how our model results change as we increase the parameter ϕ.

	Data	Model			
	(1)	(2) BTZ	(3) $\phi = 0.3$	(4) $\phi = 0.5$	(5) $\phi = 0.8$
		Risk Return Trade-Off			
$\hat{\sigma}^2_{t-1}$	−2.21	10	−0.23	−1.52	−1.05
	(1.38)				
$\Sigma_{j=1}^{6}\phi^j\hat{\sigma}^2_{t-j}$	2.17	0	3.02	3.74	1.78
	(1.42)				
R^2	0.1%	7.9%	3.5%	3.4%	1.4%
		Forecasting Returns with Variance Risk Premium			
VRP_{t-1}	3.47	3.94	7.40	2.58	1.42
	(0.92)				
R^2	2.3%	1.2%	1.9%	1%	0.5%
		Forecasting Returns with E[RV] and VIX			
$\hat{\sigma}^2_{t-1}$	−3.44	6.06	−2.28	−2.21	−2.02
	(0.87)				
VIX^2_{t-1}	3.69	3.94	4.78	4.71	4.52
	(1.25)				
R^2	2.1%	8%	4%	3.6%	3.4%
		Expected Variance Risk Premium ($VRP_{t-1} = VIX^2_{t-1} - E_{t-1}[\sigma^2_t]$)			
$\hat{\sigma}^2_{t-1}$	−1.26	0	−0.25	−0.70	−0.92
	(0.49)				
$\Sigma_{j=1}^{6}\phi^j\hat{\sigma}^2_{t-j}$	1.23	0	0.39	0.80	0.58
	(0.46)				
		Correlation: VIX^2 and Realized Variance			
	0.86	0.99	0.96	0.89	0.73

differs in terms of asset price dynamics, Table V compares our model directly to BTZ based on the stylized facts we target.[20]

Both models indicate that the variance risk premium positively predicts stock returns, which is the main fact that BTZ is calibrated to match. However, the models strongly differ on the risk-return trade-off. First, the BTZ model implies a coefficient of 10 for the risk-return trade-off, while in the data, this is

[20] To give the BTZ model a better chance at matching these patterns, we recalibrate the objective variance process in their model to match that in the data. Their calibration implies a counterfactually high persistence of RV_t.

much weaker. As a result, their model cannot match the return predictability regressions on past variance or generate a positive volatility-managed alpha. Next, a salient fact in the data is that while the difference $VIX_t^2 - RV_t$ is a strong predictor of market returns, neither lagged RV nor the VIX (or the VIX^2) are strong return predictors on their own—a fact that BTZ document in their table 3 (page 4482 in BTZ). In the BTZ model, when including both VIX_t^2 and RV_t to forecast stock returns, both coefficients are strongly positive, while in the data and our model, the coefficients have opposite signs.

E. Impulse Responses: Data versus Model

To succinctly summarize the empirical patterns in a way that is easy to relate to our model, we estimate a first-order vector autoregression (VAR) with expected variance $\hat{\sigma}^2$ (formed as before), the variance risk premium $VIX^2 - \hat{\sigma}^2$, excess market returns, and the log market price-dividend ratio as the variables in the state vector.[21] Expected variance is ordered first, so all variables can respond contemporaneously to a shock to $\hat{\sigma}^2$. Figure 2 plots impulse responses of the equity premium and variance risk premium to a shock to $\hat{\sigma}^2$.

The variance risk premium goes negative after the shock before slowly rising and becoming positive beyond month 3 (Cheng (2018)). Similarly, the equity risk premium, if anything, initially falls but rises further out. While the initial response of the equity risk premium is not statistically different from zero here (consistent with a weak or negative risk-return trade-off), note that the upper bound is small, which we show relative to leading rational models. Most notably, the premiums have a hump-shaped pattern: they appear low initially but rise as future volatility falls. This is in contrast to the standard benchmark model, with the equity premium being affine in expected variance. In this setting, the risk premium response should peak immediately and roughly mirror the response of future variance from period 1 onward, with a spike upward followed by a decline as future variance mean reverts. The shaded regions indicate 95% confidence intervals based on bootstrapping the residuals in the VAR. We consider alternative specifications of the VAR in the Internet Appendix that rely less on high-variance periods and argue that the main patterns documented here are robust.

Figure 2 also shows impulse responses in the model versus the data. The impulse responses from the calibrated model are given in the dashed blue lines, while the impulse responses from the model assuming $\phi = 0$ and $\lambda = \rho$ (the rational case) are given in the dash-dotted red lines for comparison.

The impact of a shock to expected variance on the conditional market and variance risk premiums is very different across the two models. In particular, in the rational model, the response is to immediately increase the conditional risk premium due to the usual risk-return trade-off (equation (15)), at

[21] The log price-dividend ratio is from CRSP based on value-weighted returns. We construct the price-dividend ratio as the sum of dividends over the past year divided by the current price. We find similar results using other price measures, for example, the cyclically adjusted price to earnings ratio (CAPE) from Robert Shiller's website.

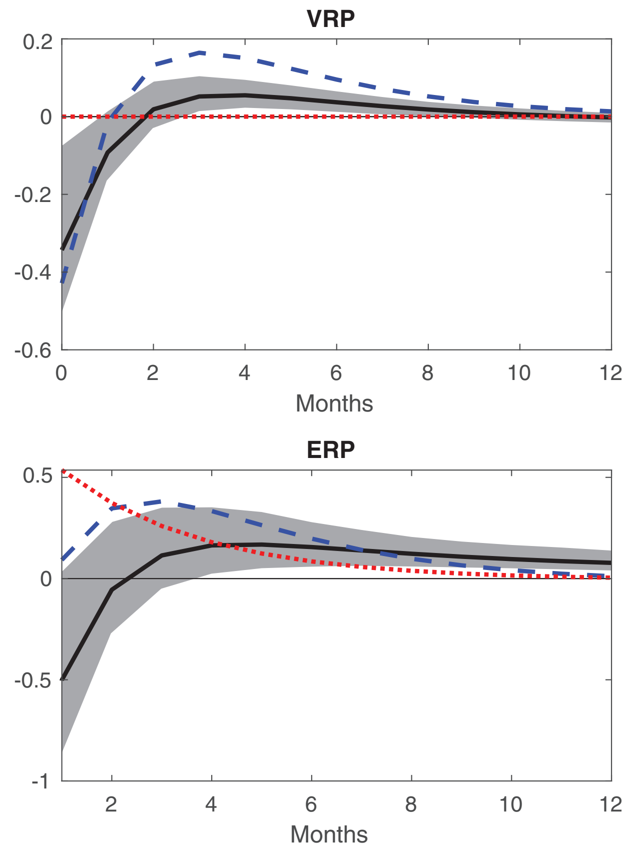

Figure 2. Impulse responses: Data versus model. We plot the behavior of expected stock returns and variance risk premiums in the data versus the model at various horizons for a one-standard-deviation shock to variance. The black line shows the impulse response in the data from a VAR(1) of expected variance, market excess returns (denoted as ERP for equity risk premium), the variance risk premium (VRP), and the log price dividend ratio. VRP is implied variance (VIX^2) minus expected variance. Responses are for a one-standard-deviation shock to expected variance at time 0 and the gray-shaded region represents 95% confidence interval. The dashed blue line repeats this using simulated data from the calibrated model. The dot-dashed red line repeats this exercise in the simulated model data but imposes no bias in beliefs (rational model). Equity returns are given in units of percent per month. The x-axis is in months. (Color figure can be viewed at wileyonlinelibrary.com)

odds with the empirical facts, whereas in our model, the response of the conditional equity premium as measured in the VAR is initially negative, as in the data. This is due to the mistake investors are making in their variance forecast as shown in Figure 1. The equity premium subsequently overshoots due to the slow-moving expectations of the agents, consistent with the pattern in the data. The same is true for the variance risk premium, although in this case, the pattern is slightly stronger than in the data as its dynamics are affected

Table VI
Stock-Level Analysis

We repeat our analysis at the stock level. We run three forecasting regressions $y_{i,t} = a_i + b\Delta_6\sigma^2_{t-1} + \varepsilon_{i,t}$, where $\Delta_6\sigma^2_{i,t-1}$ is the (lagged) six-month change in realized variance at the stock level for firm i (the realized variance estimates on the right-hand size are winsorized at the 95% level; see text for discussion). As dependent variables, y, we use the equity risk premium (stock return over the risk-free rate, $r_{i,t} - r^f_t$, labeled ERP), future variance ($\sigma^2_{i,t}$), and the variance risk premium (difference between implied variance from option metrics and future realized variance, $VRP = IV^2_{i,t-1} - \sigma^2_{i,t}$, where IV is option implied volatility). Data are monthly but realized variance uses daily data within the month. The last three columns repeat the regression using time fixed effects. In our panel regressions, standard errors are double-clustered by stock and time.

	ERP	Vol	VRP	ERP	Vol	VRP
$\Delta_6\sigma^2_{t-1}$	−0.129	0.253***	−0.104***	−0.040	0.188***	−0.070***
	(0.137)	(0.067)	(0.036)	(0.075)	(0.046)	(0.022)
N	536,726	536,726	536,726	536,726	536,726	536,726
Adj R^2	0.001	0.010	0.002	0.159	0.060	0.024
Time FE	N	N	N	Y	Y	Y

only by the mistake in expectations (see equation (22)). The rational version of the model has no effect on the variance risk premium from a shock to expected variance, again at odds with the data. While one could change this, for example, in a model where the volatility of volatility shocks varies over time, it would be difficult to generate both the negative initial response and hump-shaped dynamics. In particular, because the volatility of volatility shocks tends to be positively correlated with the level of volatility, this would tend to push the initial response of the variance risk premium up. In other words, volatility claims tend to be riskier after a shock to expected variance (Cheng (2018)), and thus, a rational model with this feature would have a harder time matching the data.

IV. Additional Evidence and Alternative Explanations

A. Firm-Level Analysis

We revisit our stylized aggregate facts at the firm level (stock level) in Table VI. We take implied volatility from OptionMetrics at the stock level from 1996 to 2017 and use daily and monthly return data from CRSP for the stocks in the merged OptionMetrics sample (6,489 unique stocks over the sample). Implied volatility is measured on the last day of the month and captures option-implied volatility over the subsequent month (30 days) for at-the-money options. Realized variance is computed using the daily returns within a given month. Our measure of the realized variance risk premium is then $IV^2_{i,t} - RV_{t+1}$, which is the implied variance over the next month minus the actual realized variance over the next month.

Similar to our results in Table III, we forecast excess equity returns, realized variance risk premiums, and future realized variance over the next month. Different from the earlier regressions, we use the change in realized variance from month t to $t-6$ as the forecasting variable, for several reasons. Most importantly, this helps account for quarterly earnings announcements at the firm level, which drive predictable spikes in firm-level volatility. By differencing the six-month lag, we account both for effects of quarterly earnings announcements on firm-level volatility and unconditional firm-level fixed effects in volatility. We winsorize the lagged change realized variance at the 95^{th} percentile, though importantly, we do not winsorize future realized variance so that the left-hand side is still the realized variance risk premium. In the Internet Appendix, we find a similar result without winsorization, but we find much stronger predictive power for future variance with winsorization due to substantially more noise in firm-level realized volatility estimates compared to the aggregate. We also find qualitatively similar results in several other specifications, including when we use log of realized variance or volatility in the place of variance, although these results are omitted for space.

The results show that increases in volatility over six months negatively forecast variance risk premiums but positively forecast future variance. The coefficients for predicting future variance and future variance risk premiums are highly statistically significant with or without time fixed effects (standard errors are double-clustered by time and firm). The results with time fixed effects are especially important because these remove any aggregate movements in firm-level variance or variance risk premiums.

The firm-level analysis makes two contributions. First, it provides robustness to our aggregate results, which rely on fewer observations. Second, it provides more insight into whether the variance risk premium results that we document are likely driven by true economic risk premiums (compensation for risk) or by biased expectations and underreaction to changes in volatility. As we stress earlier, the aggregate results are not consistent with standard risk-based models since higher risk (more variance) should, if anything, imply a higher rather than lower risk premium. Nevertheless, it is always possible to construct a model in which investor preferences move in such a way to match the aggregate evidence. The firm-level evidence is more powerful since we think of firm-level variance as largely idiosyncratic, especially in our second specification where we include time fixed effects in the regression to remove any common components of firm-level variance. Hence, we might expect a much smaller effect at the firm level from a risk premium story due to variance shocks being more idiosyncratic at the firm level. Instead, we recover a coefficient of around -0.1 for the firm-level variance risk premium, which is in line with the magnitudes we observe in the aggregate results. These firm-level results are also similar in spirit to Poteshman (2001), who argues for underreaction in option prices in an earlier sample.

Further in this dimension, at the firm level, we see a weakly negative but not significant coefficient on the equity risk premium. This is exactly what we expect in the model if agents do not price idiosyncratic firm-level risk. In our aggregate results, investors should require more compensation for the increase in variance, and this mechanism combined with biased beliefs results in the negative coefficient on the equity risk premium. Absent this channel, we would only expect the results to hold for the variance risk premium. Taken together, the firm-level results support our main hypothesis that agents initially underreact to changes in variance and that this is reflected in implied volatilities.

B. Evidence on Actual Trading Behavior

Hoopes et al. (2021) provide evidence that investors do react to changes in volatility, with more sophisticated investors and older investors responding more strongly. Specifically, they show that higher income and older investors sell more aggressively following increases in volatility. This is reasonable in our model if one takes higher income investors to be more sophisticated and less prone to the expectations bias in our paper. Similarly, it is possible that investors learn more about the volatility process over time (as the evidence on investor experience suggests they would) and hence exhibit less of a bias as they become older. Giglio et al. (2019) also find that agents' views about risk are informative about their portfolio decisions, with higher expectations of stock market risk being associated with a lower allocation to stocks. A shortcoming of our model is that it features a representative investor and hence does not speak directly to this evidence (as there is no trade in equilibrium), though modest extensions of the model that allow for differences in the amount of bias would naturally be consistent with the evidence on trading behavior in Hoopes et al. (2021).

C. Term Structure Evidence

The model has implications for the conditional term structure of variance risk premiums. Above we show that the model produces underreaction to volatility for one-month variance claims. A natural question is how this extends to longer horizon variance claims.

Empirically, we have return data for variance swaps, VIX futures, and straddles for multiple maturities. We show these implications in Table VII where we run the same predictive regressions as before using all maturities and normalize the coefficients by dividing by 100. For VIX futures and straddles, we focus on the sample in which we have nonmissing observations at all maturities using data provided by Johnson (2017). We run the analogous procedure in the model where we regress the one-month return of a variance swap with maturity k on expected variance and a weighted average of expected variance. Overall, current expected variance continues to negatively predict returns for all maturities, while past variance positively predicts. The slow-moving nature of the extrapolative expectations means that agents believe that variance

Table VII

Term Structure

We regress excess returns to variance claims at different maturities on lagged variance and a weighted average of past variance. Data and returns are monthly, and standard errors in parentheses are Newey-West with 12 lags. In the data, coefficients are normalized for comparison to the model. See the text for more detail.

Panel A: Model Coefficients by Maturity

	(1)	(2)	(3)	(6)	(12)
σ_{t-1}^2	−0.34	−0.20	−0.15	−0.14	−0.12
$\Sigma_{j=1}^6 \phi^j \hat{\sigma}_{t-j}^2$	0.54	0.31	0.24	0.17	0.14
Adj. R^2	15.9%	15.9%	15.9%	15.9%	15.9%

Panel B: Data Straddle Returns

	(1)	(2)	(3)	(6)	(12)
σ_{t-1}^2	−0.29**	−0.20***	−0.14***	−0.10***	−0.08***
	(0.11)	(0.04)	(0.04)	(0.02)	(0.01)
$\Sigma_{j=1}^6 \phi^j \hat{\sigma}_{t-j}^2$	0.37**	0.25***	0.18***	0.14***	0.11***
	(0.14)	(0.05)	(0.05)	(0.03)	(0.02)
N	292	282	282	282	282
Adj. R^2	2.4%	3.1%	2.4%	2.9%	3.5%

Panel C: Data Variance Swaps

	(1)	(3)	(6)	(12)
σ_{t-1}^2	−0.41***	−0.28***	−0.18***	−0.08**
	(0.14)	(0.08)	(0.05)	(0.04)
$\Sigma_{j=1}^6 \phi^j \hat{\sigma}_{t-j}^2$	0.43**	0.29***	0.17**	0.06
	(0.20)	(0.11)	(0.07)	(0.05)
N	264	264	264	264
Adj. R^2	0.7%	1.4%	1.3%	0.0%

Panel D: Data VIX Futures

	(1)	(3)	(6)
σ_{t-1}^2	−0.35***	−0.33***	−0.20***
	(0.08)	(0.06)	(0.04)
$\Sigma_{j=1}^6 \phi^j \hat{\sigma}_{t-j}^2$	0.26**	0.03	−0.17*
	(0.13)	(0.09)	(0.09)
N	166	166	166
Adj. R^2	2.2%	3.6%	5.1%

is highly persistent, which, in turn, implies that mistakes in conditional variance expectations also matter for long-horizon claims. This generates excess volatility in the long end of the variance term structure, matching the spirit of the empirical findings in Giglio and Kelly (2017).

Table VIII
Additional Returns

We run predictive regressions of future excess returns on past expected variance, and a weighted average of expected variance over the past six months, $(1-\phi)\Sigma_{k=1}^{6}\phi^{k-1}\hat{\sigma}_{t-k}^{2}$. We use credit returns as credit is a volatility-sensitive asset class. We use credit default swap (CDS) returns from He, Kelly, and Manela (2017), the difference in Barclays high-yield and investment-grade returns (HY-IG), and the (negative) change in the Baa-Aaa spread from Moody's. Data are monthly from 1990 to 2018, the CDS data and high-yield total return data are from 2004 to 2012 and 1995 to 2015, respectively. Standard errors in parentheses use Newey-West correction with 12 lags.

	Credit Returns		
	CDS (1)	HY-IG (2)	Baa-Aaa (3)
σ_{t-1}^{2}	−1.21***	−5.74***	−0.34***
	(0.13)	(1.71)	(0.06)
$\Sigma_{j=1}^{6}\phi^{j}\hat{\sigma}_{t-j}^{2}$	1.32***	5.28**	0.33***
	(0.20)	(2.30)	(0.08)
N	143	240	313
Adj. R^2	19.2%	15.9%	25.9%

D. Credit Returns

Table VIII again runs predictive regressions but uses returns that depend on credit risk. These returns are particularly interesting for our story because credit risk is sensitive to changes in volatility that affect default risk (Merton (1974)), and thus constitute an additional robustness check for our main results. We construct credit returns using three sources. The first comprises CDS excess returns from He, Kelly, and Manela (2017), who form 20 CDS portfolios based on credit risk. We equally weight across these 20 portfolios to form a single CDS excess return. In the second column, we use the Barclays total return index for high-yield and investment-grade corporate bonds. We compute the excess return as the difference between the return of these two series. The third source is the (negative) change in the Baa-Aaa yield spread from Moody's. This proxy for the credit return is previously used in López-Salido, Stein, and Zakrajšek (2017).

All three credit return series point in the same direction and support our main results. Current volatility negatively predicts returns where the moving average of past variance positively predicts returns. Similar to our earlier regressions, the absolute magnitudes of the coefficients are similar, meaning that the change in volatility relative to the moving average explains the predictability. Both coefficients are highly significant in all three cases.

E. Model Shortcomings and Extensions

Our model is simplified to focus on the role of one particular channel – beliefs about volatility – influencing the prices of financial market claims. Below we outline limitations of the baseline model and discuss useful extensions.

Our model has implications for the price-dividend ratio that are clearly rejected in the data. Most importantly, if taken literally, the model says that the dividend yield is perfectly correlated with the VIX in levels, which is clearly counterfactual. Empirically, the dividend yield is much more persistent than the VIX, although the two are correlated (the VIX and the dividend yield both tend to go up in bad times). This suggests that the dividend yield is likely also influenced by forces outside our model. An extension of our model with time-varying expected dividend growth would generate additional movements in the dividend yield and, if these growth rates were highly persistent, could generate the difference in persistence and also reduce the correlation between the dividend yield and the VIX in levels. A robust stylized fact is that changes in prices are highly negatively correlated with changes in the VIX (prices go down when the VIX goes up) as in our model, and this result can still hold in an extension with persistent time variation in expected dividend growth. This fact highlights why we choose not to use dividend yield-related moments when targeting the parameters in our calibration even though our baseline model has implications for these moments.

In the main model for convenience, we put the stochastic volatility on the cash flow process. However, this is not particularly important and the main mechanism we advocate is that agents make mistakes in return variance forecasting, whether that is discount rate volatility or cash flow volatility.

In our model, RV follows an AR(1) process. Thus, the optimal predictor of variance in the model is simply lagged RV, while the VIX has no marginal predictive power. In the sample we consider (1990 to 2020), this is, in fact, a close approximation—the increase in R^2 in a forecasting regression for RV is negligible when adding the VIX as a predictive variable in addition to lagged RV. In other samples, however, the VIX has stronger marginal predictive power. Chernov (2007) shows that this is indeed the case in the 1986 to 2001 sample, but he also points out that the VIX alone cannot be the optimal predictor of future variance since the variance risk premium is time-varying in the data.[22] This latter fact is consistent with our model, where the VIX reflects agents' expectational errors in addition to the objective variance forecast. In terms of the latter, it is straightforward to extend our model to make the VIX have marginal predictive power for future variance. In particular, if we assume that each period investors observe a noisy signal about next period's RV that is uncorrelated with current RV (i.e., a time t signal correlated with η_{t+1}), then the VIX will reflect this added information as the signal affects agents' beliefs. Importantly, such an extension will not affect the projection of the VIX or agents' beliefs about future variance onto lagged RVs, which is the focus of the bias we consider in our model. We provide this extension in Section III of the Internet Appendix.

[22] Relatedly, Chernov (2007) finds that the coefficient on the VIX in variance forecasting regressions is significantly below one, consistent with substantial variation in the variance risk premium. This is also the case in our sample.

A natural extension of our model is to introduce a set of agents (risk-averse arbitrageurs) that have rational expectations. In such a model, the wealth share of the arbitrageurs becomes a new state variable. Generally speaking, the degree of mispricing will increase with impediments to arbitrage and decrease with the arbitrageur wealth share. Several features of the data are consistent with this. For instance, we document a stronger pattern when current variance is high. Since arbitrageurs on average are short volatility, these are times when they have recently suffered losses on their arbitrage positions and therefore hold less wealth, scaling back their positions. This is consistent with data on hedge fund positions as shown by Cheng (2018). Further, our results are strong after 2010, which coincides with a period of tighter bank regulation after the financial crisis. Also, our results showing that a higher ϕ is needed to match the moments in the stock market than the moments in the variance market is consistent, as there are many shocks to stock prices beyond variance shocks, making the arbitrage riskier in the stock market. While an extension to such a heterogeneous-agent model would be quite interesting, it is beyond the scope of this paper.

One implication and limitation of our representative agent model is that expectations of returns and variance should be strongly positively related. However, empirically survey expectations of returns and variance from the Graham and Harvey CFO survey are essentially uncorrelated: the adjusted R^2 of survey expected returns regressed on our measure of expected variance is zero. Giglio et al. (2019) show at the *individual* level, expectations returns and expectations of a crash in the stock market are strongly *negatively* correlated. These results are natural if agents have heterogeneous beliefs, which our representative agent model does not capture. For example, as Giglio et al. (2019) mention, if an individual investor assigns a higher probability to a market crash than do other agents, that investor will think that future risk is high and that expected returns are low. The reason is that this agent is small relative to the market, so their beliefs affect market prices little if at all. The agent will thus view stock prices as too high and future returns as too low since they do not reflect the risk of a crash, and as a result will decrease their portfolio allocation in their data as expected. This would not be the case if this were the only investor in the market, as prices would then need to fall substantially to keep the agent from selling.

However, as we aggregate agents' views, the correlation between expectations of returns and expectations of risk will increase. If a large subset of investors believes that the risk of a crash is high, this view will be reflected in aggregate market prices, and thus, raise their view of expected returns. Our CFO survey results are consistent with this story: the correlation between aggregated CFO views on variance and expected returns in our survey evidence is much higher than the individual level results of Giglio et al. (2019), although are still not nearly as high as in our model with a single investor. There are typically around 300 responses to the CFO survey (Graham and Harvey (2008)), which is far from the aggregate expectations of the entire market. This view is also related to Greenwood and Shleifer (2014), who acknowledge that not all investors can have low expectations of returns or be return extrapolators, – there must be other investors on the other side that are absent in these surveys.

We view this heterogeneity in beliefs as important, and believe that extensions of our model that account for these facts at the individual level would be fruitful areas to study, as well as extensions in which some agents extrapolate risk and some have rational beliefs. Because accommodating these features in our current framework would add a substantial amount of complexity, we leave this to future research.

F. Rational Arbitrageurs and Trading Profits

The biased beliefs of the representative agent in our model lead to errors in expectations about conditional stock market variance. In this section, we analyze the gains a rational agent would achieve through optimal timing of the variance claim, assuming that this agent trades only an infinitesimally small amount and thus does not affect prices. We focus on the variance claim because this is where trading on the mistake is most direct and has the highest Sharpe ratio (as compared to the equity claim).

From equation (22), we have that variation in the variance risk premium is due to differences between the agent's subjective belief about conditional stock market variance and the true conditional stock market variance. Consider a myopic mean-variance optimizing rational agent that is timing the variance claim based on the current expectational errors of the representative agent. The rational agent's position in the variance claim is then

$$\omega_t = \frac{VRP_t}{\gamma \sigma_\eta^2}, \tag{26}$$

where γ is the coefficient of risk-aversion and ω_t is the number of units short in the variance claim. That is, when the expected return to shorting variance is high (low), the short variance claim position is scaled up (down). To focus on the effects of timing, we consider the Sharpe ratio of a strategy that is on average variance neutral. That is, we compute the Sharpe ratio for the strategy with returns $(\omega_t - E[\omega_t])(IV_t - RV_{t+1})$ both in the model and in the data, noting that the risk aversion parameter γ does not need to be specified for the Sharpe ratio calculation.

To estimate the VRP in the data, we run the regression

$$IV_{t-1} - RV_t = \alpha + \beta_1 \hat{\sigma}^2_{t-1} + \beta_2 IV_{t-1} + \eta_t, \tag{27}$$

where IV is the squared VIX and $\hat{\sigma}^2$ is the heterogeneous autoregressive estimate (HAR) of the true conditional market variance. The estimated VRP is then $\widehat{VRP}_t = \hat{\alpha} + \hat{\beta}_2 \hat{\sigma}^2_t + \hat{\beta}_2 IV_t$. This is the optimal regression to run within the model for estimating the conditional variance risk premium. Table IX presents the results of this regression both in the data and in the model. The first column shows that the $\hat{\beta}_1$ and $\hat{\beta}_2$ coefficients in the data are −0.85 and 0.87, while

Table IX

Timing Sharpe Ratios

We run predictive regressions of the realized variance risk premium ($VIX_{t-1}^2 - RV_t$) on expected variance and the squared VIX at $t-1$. We then take the predicted value from the regression and compute the Sharpe ratio of the optimal timing portfolio. The first column is the data and the second two columns are the main model and our extended model with jumps and time-varying volatility of volatility. Standard errors in parentheses use Newey-West correction with 12 lags.

	Data (1)	Model (2)	Model Extended (3)
$\hat{\sigma}_{t-1}^2$	-0.85^{***}	-1	-1
	(0.20)		
VIX_{t-1}^2	0.87^{***}	1	1
	(0.18)		
N	338		
R^2	0.19	0.124	0.09
Sharpe	0.41	1.17	0.30

the R^2 is 19%. We note that the regression coefficients are not statistically different from -1 and 1, respectively. The annualized sample Sharpe ratio from the above timing strategy is 0.41.

In column (2), we run the same regression on simulated model data. The regression coefficients in the model are indeed -1 and 1, while the R^2 is 12%, slightly lower than the regression in the data. The timing Sharpe ratio in the model, however, is quite a bit larger than that in the data at 1.17.

The results suggest that the mistakes in variance expectations within our model are too large. However, the high timing Sharpe ratio is not a robust feature of our model. In particular, in the baseline model, we assume for analytical convenience assume that conditional variance is Normally distributed, but realized variance in the data has sample skewness and kurtosis of 6.89 and 69, respectively. To illustrate the effect of a more realistic data process for variance without altering the expectation formation mechanism, we simulate variance as

$$\sigma_t^2 = \bar{v} + \rho\left(\sigma_{t-1}^2 - \bar{v}\right) + \omega\sqrt{\sigma_{t-1}^2}\tilde{\eta}_t, \tag{28}$$

where $\tilde{\eta}_t = \eta_t^{(1)} + \eta_t^{(2)} J_t$ is a mixture of Normals. In particular, $\eta_t^{(1)} \sim N(0, 1)$, $\eta_t^{(2)} \sim N(0, \sigma_J^2)$, and J_t is a random indicator variable that each period equals one with probability p and zero otherwise. Thus, we have both time-varying volatility and "jumps" in the variance process that together with the square root process allow us to achieve high skewness and kurtosis. Eraker, Johannes, and Polson (2003) show that such dynamics are important to account for the dynamic behavior of conditional market return variance. In continuous time, this process never goes negative and in our discrete-time simulations, we simply set any negative variances to zero, although we note that such observations are rare and small in magnitude. We calibrate this process to match the first four moments of realized variance by setting $\omega = 0.0738$, $\sigma_J = 5$, and $p = 1/30$ and using the same persistence parameter as before, $\rho = 0.71$. Under the assumption that variance risk is not priced and with the same process for x_t as before, the impulse response of the variance risk premium to a variance shock is the same as in the baseline model at all horizons. In other words, the main features of our model remain the same in this simulation.

The third column of Table IX shows that this more realistic specification of variance dynamics again gives regression coefficients of -1 and 1 as in the baseline case. The R^2 is 9% and the Sharpe ratio from timing is now much smaller at 0.30, slightly lower than in the data. We note that an annualized Sharpe ratio of around 0.5 is similar to that of other anomalies such as momentum or carry.

G. Alternative Explanations

Moreira and Muir (2017) show that leading equilibrium asset pricing models (e.g., habits models, intermediary models, long-run risk models, and rare disasters models) typically imply a strong risk-return trade-off and hence do not match the fact that volatility is a weak predictor of returns.

What other models could explain our results? While some models can indeed match some of our stylized facts, we are not aware of models that can jointly match them quantitatively. This is especially true for the firm-level analysis, which relies solely on idiosyncratic movement in firm-level variance, and our survey expectation data, which suggest slow-moving volatility expectations. We briefly discuss models with rational inattention, heterogeneity, and learning in terms of which facts they can explain.

Models featuring infrequent rebalancing and/or rational inattention (Abel, Eberly, and Panageas (2013)) at first appear promising but do not easily match the facts that we document. Even if agents know that they will not rebalance again for a while, due to rational inattention, they will still ensure a risk-return trade-off at the horizon at which they expect to rebalance. This will result in a risk-return trade-off that resembles the standard case. Further, Hoopes et al. (2021) show evidence that investors do react to changes in volatility with more sophisticated investors (e.g., those in highest income brackets) responding most quickly. That is, it does not appear that agents are not aware and do not act on changes in volatility. Finally, we are unaware of these models being able to easily match the variance risk premium dynamics, particularly the firm-level facts or the survey expectation data.

Heterogeneous-agent models can potentially explain the weak risk-return relation, and in these models, this relation can even go negative depending on the wealth distribution (e.g., Longstaff and Wang (2012), Gârleanu and Panageas (2015)). These models feature a conditional risk-return trade-off that is typically positive for most parts of the stationary distribution but can turn negative in the worst states. For the unconditional risk-return trade-off to be weak, calibrations of the models would typically also require that the correlation between shocks to returns and volatility would be weak. Empirically, this correlation is strongly negative, and this stylized fact is robust across good and bad market conditions. Further, in typical calibrations that aim to match other moments (e.g., the equity premium), the risk-return trade-off is positive. It is not obvious that these models would be able to explain the mismatch in frequencies that we observe, for example, with risk premiums initially declining but then rising further out after volatility increases, the variance risk premium results, the firm-level results (which rely on firm-level idiosyncratic variance rather than aggregate variance), or the slow-moving expectations from our survey data.

Since high current volatility reflects high uncertainty about asset values, it is natural to consider learning about parameters and/or latent states governing economic dynamics as a possible explanation for the empirical facts. Consider as an example a model in which there are two latent high-volatility regimes—one regime is short-lived, the other regime is long-lived. Let us say current volatility is high, so investors know that the economy is in one of these high-volatility regimes but they do not know which one. In a stationary learning environment with a long sample, investors' prediction for future volatility will on average be right given available information, so there should not be return predictability in returns other than that coming from higher expected risk. However, the learning may be nonstationary and/or the sample may be relatively small when it comes to these types of events. Thus, ex post predictability that was not ex ante actionable may appear in forecasting regressions. For instance, at the onset of the financial crises, anecdotally, there was substantial fear of a long-lived crisis as volatility spiked to extreme levels. In fact, volatility instead mean-reverted quite quickly as investors learned that the financial meltdown scenario was averted. Note that in this case, investors would have overestimated future variance, as the high-volatility regime ex post turned out to be of the short-lived variety. Thus, investors would appear to have overreacted to the initial high-volatility shock, which would lead to high variance forecasting high returns. This is the opposite of what we find. That said, the concern of small-sample issues is valid, which is one of the reasons we also examine the cross-section of firm-level options. The consistent results in a pure cross-sectional analysis suggest that our results are not due to small-sample concerns.

Some learning models can explain a weak or even negative relation between prices, expected returns, and conditional return variance. Examples include Veronesi (2000) and David and Veronesi (2013). For instance, high uncertainty about the growth rate could lead to high prices through a convexity effect along with high return variance. However, in the data, shocks to market volatility are robustly negatively correlated with contemporaneous market returns. Further, these models do not also address the joint evidence on return predictability in stock and variance claim markets, or the survey evidence on volatility expectations.

V. Conclusion

We show that underreaction followed by delayed overreaction to volatility news can match many empirical facts surrounding volatility and risk premiums that are puzzling from the perspective of leading equilibrium asset pricing models. These results obtain by assuming that agents' expectations of volatility are slow moving and extrapolative, and are consistent with expectations in survey data. In particular, our model matches the weak overall risk-return trade-off and the dynamic responses of both the equity premium and the variance risk premium following shocks to variance.

We are able to account for the fact that shocks to volatility are indeed associated with negative contemporaneous realized returns through a discount rate channel, although the relation between volatility and next-period returns is weak. Finally, in our model, the variance risk premium predicts returns more strongly than either variance or implied variance, as in the data. Survey evidence directly supports slow-moving expectations about volatility, as does evidence using firm-level option prices.

Initial submission: December 30, 2019; Accepted: February 26, 2021
Editors: Stefan Nagel, Philip Bond, Amit Seru, and Wei Xiong

REFERENCES

Abel, Andrew B., Janice C. Eberly, and Stavros Panageas, 2013, Optimal inattention to the stock market with information costs and transactions costs, *Econometrica* 81, 1455–1481.

Bansal, Ravi, Dana Kiku, and Amir Yaron, 2016, Risks for the long run: Estimation with time aggregation, *Journal of Monetary Economics* 82, 52–69.

Bansal, Ravi, and Amir Yaron, 2004, Risks for the long run: A potential resolution of asset pricing puzzles, *Journal of Finance*, 1481–1509.

Barberis, Nicholas, Robin Greenwood, Lawrence Jin, and Andrei Shleifer, 2015, X-CAPM: An extrapolative capital asset pricing model, *Journal of Financial Economics* 115, 1–24.

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny, 1998, A model of investor sentiment, *Journal of Financial Economics* 49, 307–343.

Barro, Robert J., 2006, Rare disasters and asset markets in the twentieth century, *Quarterly Journal of Economics* 121, 823–866.

Bekaert, Geert, and Marie Hoerova, 2014, The VIX, the variance premium and stock market volatility, *Journal of Econometrics* 183, 181–192.

Ben-David, Itzhak, John R. Graham, and Campbell R. Harvey, 2013, Managerial miscalibration, *Quarterly Journal of Economics* 128, 1547–1584.

Bollerslev, Tim, James Marrone, Lai Xu, and Hao Zhou, 2014, Stock return predictability and variance risk premia: Statistical inference and international evidence, *Journal of Financial and Quantitative Analysis* 49, 633–661.

Bollerslev, Tim, George Tauchen, and Hao Zhou, 2009, Expected stock returns and variance risk premia, *Review of Financial Studies* 22, 4463–4492.

Bollerslev, Tim, Viktor Todorov, and Lai Xu, 2015, Tail risk premia and return predictability, *Journal of Financial Economics* 118, 113–134.

Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer, 2020, Overreaction in macroeconomic expectations, *American Economic Review* 110, 2748–82.

Bouchaud, Jean-philippe, Philipp Krueger, Augustin Landier, and David Thesmar, 2019, Sticky expectations and the profitability anomaly, *Journal of Finance* 74, 639–674.

Boutros, Michael, Itzhak Ben-David, John R. Graham, Campbell R. Harvey, and John W. Payne, 2019, The persistence of miscalibration. Available at SSRN 3462107.

Brandt, Michael W., and Qiang Kang, 2004, On the relationship between the conditional mean and volatility of stock returns: A latent VAR approach, *Journal of Financial Economics* 72, 217–257.

Brooks, Jordan, Michael Katz, and Hanno Lustig, 2018, Post-FOMC announcement drift in U.S. bond markets. Technical report, National Bureau of Economic Research.

Campbell, John Y., and John Cochrane, 1999, By force of habit: A consumption-based explanation of aggregate stock market behavior, *Journal of Political Economy* 107, 205–251.

Campbell, John Y., and Robert J. Shiller, 1988, The dividend-price ratio and expectations of future dividends and discount factors, *Review of Financial Studies* 1, 195–228.

Cheng, Ing-Haw, 2018, The VIX Premium, *Review of Financial Studies* 32, 180–227.

Chernov, Mikhail, 2007, On the role of risk premia in volatility forecasting, *Journal of Business & Economic Statistics* 25, 411–426.

Coibion, Olivier, and Yuriy Gorodnichenko, 2015, Information rigidity and the expectations formation process: A simple framework and new facts, *American Economic Review* 105, 2644–2678.

Collin-Dufresne, Pierre, Michael Johannes, and Lars A. Lochstoer, 2016, Asset pricing when 'this time is different', *Review of Financial Studies* 30, 505–535.

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam, 1998, Investor psychology and security market under- and overreactions, *Journal of Finance* 53, 1839–1885.

David, Alexander, and Pietro Veronesi, 2013, What ties return volatilities to price valuations and fundamentals? *Journal of Political Economy* 121, 682–746.

Dew-Becker, Ian, Stefano Giglio, Anh Le, and Marius Rodriguez, 2017, The price of variance risk, *Journal of Financial Economics* 123, 225–250.

Drechsler, Itamar, and Amir Yaron, 2011, What's Vol got to do with it, *Review of Financial Studies* 24, 1–45.

Epstein, Larry G., and Stanley E. Zin, 1989, Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework, *Econometrica: Journal of the Econometric Society* 57, 937–969.

Eraker, Bjorn, 2020, Predictability puzzles. Working Paper.

Eraker, Bjørn, Michael Johannes, and Nicholas Polson, 2003, The impact of jumps in volatility and returns, *Journal of Finance* 58, 1269–1300.

French, Kenneth R., G. William Schwert, and Robert F. Stambaugh, 1987, Expected stock returns and volatility, *Journal of Financial Economics* 19, 3–29.

Fuster, Andreas, Ben Hebert, and David Laibson, 2011, Natural expectations, macroeconomic dynamics, and asset pricing, *NBER Macroeconomics Annual* 26, 1–48.

Gârleanu, Nicolae, and Stavros Panageas, 2015, Young, old, conservative, and bold: The implications of heterogeneity and finite lives for asset pricing, *Journal of Political Economy* 123, 670–685.

Gibbons, Michael R., Stephen A. Ross, and Jay Shanken, 1989, A test of the efficiency of a given portfolio, *Econometrica: Journal of the Econometric Society* 57, 1121–1152.

Giglio, Stefano, and Bryan Kelly, 2017, Excess volatility: Beyond discount rates, *Quarterly Journal of Economics* 133, 71–127.

Giglio, Stefano, Matteo Maggiori, Johannes Stroebel, and Stephen Utkus, 2019, Five facts about beliefs and portfolios. Technical report, National Bureau of Economic Research.

Glaeser, Edward L., and Charles G. Nathanson, 2017, An extrapolative model of house price dynamics, *Journal of Financial Economics* 126, 147–170.

Glosten, Lawrence R., Ravi Jagannathan, and David E. Runkle, 1993, On the relation between the expected value and the volatility of the nominal excess return on stocks, *Journal of Finance* 48, 1779–1801.

Graham, John R., and Campbell R. Harvey, 2008, The equity risk premium in 2008: Evidence from the global cfo outlook survey. Working paper available at SSRN 1162809.

Greenwood, Robin, and Andrei Shleifer, 2014, Expectations of returns and expected returns, *Review of Financial Studies* 27, 714–746.

He, Zhiguo, Bryan Kelly, and Asaf Manela, 2017, Intermediary asset pricing: New evidence from many asset classes, *Journal of Financial Economics* 126, 1–35.

He, Zhiguo, and Arvind Krishnamurthy, 2013, Intermediary asset pricing, *American Economic Review* 103, 732–770.

Hong, Harrison, and Jeremy C. Stein, 1999, A unified theory of underreaction, momentum trading, and overreaction in asset markets, *Journal of Finance* 54, 2143–2184.

Hoopes, Jeffrey L., Patrick Langetieg, Stefan Nagel, Daniel Reck, Joel Slemrod, and Bryan A. Stuart, 2021, Who sells during a crash? Evidence from tax return data on daily sales of stock, *Economic Journal* 132, 299–325.

Jiang, Zhengyang, Arvind Krishnamurthy, and Hanno Lustig, 2018, Foreign safe asset demand and the dollar exchange rate. Working paper 24439, National Bureau of Economic Research.

Johnson, Travis L., 2017, Risk premia and the VIX term structure, *Journal of Financial and Quantitative Analysis* 52, 2461–2490.

Johnson, Travis L., 2019, A fresh look at return predictability using a more efficient estimator, *Review of Asset Pricing Studies* 9, 1–46.

Landier, Augustin, Yueran Ma, and David Thesmar, 2019, Biases in expectations: Experimental evidence. Technical report, Working paper.

Lettau, Martin, and Sydney C. Ludvigson, 2010, Measuring and modeling variation in the risk-return trade-off, in *Handbook of Financial Econometrics: Tools and Techniques*, volume 1, 617–690 (North-Holland, Amsterdam, Netherlands).

Longstaff, Francis A., and Jiang Wang, 2012, Asset pricing and the credit market, *Review of Financial Studies* 25, 3169–3215.

López-Salido, David, Jeremy C. Stein, and Egon Zakrajšek, 2017, Credit-market sentiment and the business cycle, *Quarterly Journal of Economics* 132, 1373–1426.

Malmendier, Ulrike, and Stefan Nagel, 2011, Depression babies: Do macroeconomic experiences affect risk taking? *Quarterly Journal of Economics* 126, 373–416.

Mankiw, N. Gregory, and Ricardo Reis, 2002, Sticky information versus sticky prices: A proposal to replace the New Keynesian Phillips Curve*, *Quarterly Journal of Economics* 117, 1295–1328.

Martin, Ian, 2016, What is the expected return on the market? *Quarterly Journal of Economics* 132, 367–433.

Merton, Robert C., 1974, On the pricing of corporate debt: The risk structure of interest rates, *Journal of Finance* 29, 449–470.

Moreira, Alan, and Tyler Muir, 2017, Volatility-managed portfolios, *Journal of Finance* 72, 1611–1644.

Moreira, Alan, and Tyler Muir, 2019, Should long-term investors time volatility? *Journal of Financial Economics* 131, 507–527.

Nagel, Stefan, and Zhengyang Xu, 2019, Asset pricing with fading memory. Working paper, University of Chicago.

Payzan-LeNestour, Elise, Lionnel Pradier, and Tālis J. Putniņš, 2018, The waterfall illusion in the financial markets: Risk perception is distorted after prior exposure to extreme risk. Working paper, UNSW Business School.

Poteshman, Allen M., 2001, Underreaction, overreaction, and increasing misreaction to information in the options market, *Journal of Finance* 56, 851–876.

Rachwalski, Mark, and Quan Wen, 2016, Idiosyncratic risk innovations and the idiosyncratic risk-return relation, *Review of Asset Pricing Studies* 6, 303–328.

Stambaugh, Robert F., 1999, Predictive regressions, *Journal of Financial Economics* 54, 375–421.

Stein, Jeremy, 1989, Overreactions in the options market, *Journal of Finance* 44, 1011–1023.

Veronesi, Pietro, 2000, How does information quality affect stock returns? *Journal of Finance* 55, 807–837.

Wachter, Jessica A., 2013, Can time-varying risk of rare disasters explain aggregate stock market volatility? *Journal of Finance* 68, 987–1035.

Whitelaw, Robert F., 1994, Time variations and covariations in the expectation and volatility of stock market returns, *Journal of Finance* 49, 515–541.

Zhou, Hao, 2018, Variance risk premia, asset predictability puzzles, and macroeconomic uncertainty, *Annual Review of Financial Economics* 10, 481–497.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Appendix S1: Internet Appendix.
Replication Code.