How Should Investors Respond to Increases in Volatility?

Alan Moreira and Tyler Muir*

December 3, 2016

Abstract

They should reduce their equity position. We study the portfolio problem of a long-horizon investor that allocates between a risk-less and a risky asset in an environment where both volatility and expected returns are time-varying. We find that investors, regardless of their horizon, should substantially decrease risk exposure after an increase in volatility. Ignoring variation in volatility leads to large utility losses (on the order of 35% of lifetime utility). The utility benefits of volatility timing are larger than those coming from expected return timing (i.e., from return predictability) for all investment horizons we consider, particularly when parameter uncertainty is taken into account. We approximate the optimal volatility timing portfolio and find that a simple two fund strategy holds: all investors choose constant weights on a buy-and-hold portfolio and a volatility timing portfolio that scales the risky-asset exposure by the inverse of expected variance. We then show robustness to cases where the degree of mean-reversion in stock returns co-moves with volatility over time.

*Yale School of Management and UCLA Anderson School of Management. We thank John Campbell, John Cochrane, William Goetzmann, Ben Hebert, Jon Ingersoll, Ravi Jagannathan, Serhiy Kosak, Hanno Lustig, Justin Murfin, Stefan Nagel, Lubos Pastor, Myron Scholes, Ivan Shaliastovich, Ken Singleton, Tuomo Vuoltenahoo, Lu Zhang and participants at Yale SOM, UCLA Anderson, Stanford GSB, and Arrowstreet Capital for comments. We especially thank Nick Barberis for many useful discussions.

Stock market volatility is highly variable and easily forecastable, yet it is a conventional view among many practitioners and academics that investors should sit tight and not sell after increases in volatility which typically follow market downturns. Furthermore, it is often argued that *long-term* investors should view these high volatility periods as unique buying opportunities. In this paper, we investigate this conventional view. Specifically, we answer two questions: (1) how much volatility timing should investors do, if any, and (2) what are the utility benefits of volatility timing? Our approach is to study the portfolio problem of a long-lived investor that allocates her wealth between a risk-less and a risky asset in an environment where both volatility and expected returns are time-varying. We then provide comprehensive and quantitative answers to these questions and show how our answers depend on the investor's horizon and their risk aversion. Importantly, our analysis also takes into account that investors' face parameter uncertainty regarding the dynamics of volatility and expected returns.

Our main finding is that investors should substantially *decrease* risk exposure after an increase in volatility and that ignoring variation in volatility leads to large utility losses. The benefits of volatility timing are on the order of 35% of lifetime utility for our preferred parameterization of an investor with risk aversion of 5 and a 20 year horizon. These benefits are significantly larger than those coming from expected return timing (i.e., from return predictability), particularly when parameter uncertainty is taken into account. We approximate the optimal volatility timing portfolio and find that its dependence on volatility is very simple: all investors, regardless of horizon, will choose fixed weights on a buy-and-hold portfolio that invests a constant amount in the risky-asset, and a volatility managed portfolio that scales the risky-asset exposure by the inverse of expected variance $1/\sigma_t^2$. Further, we show that the weight on the volatility timing portfolio is independent of the investors' horizon in our baseline results. In contrast, the weight on the buy-and-hold portfolio depends strongly on horizon and the amount of mean reversion investors' perceive in stock returns, but doesn't depend on the dynamics of volatility. Thus, despite an apparently complex numerical exercise, our solution turns out to be simple and intuitive.

We begin our analysis by estimating a rich model for the dynamics of excess stock returns using simulated method of moments (SMM) and the last 90 years of stock return data. Our process for returns allows for time-variation in both volatility and expected returns. Allowing for both features is essential to capture the common argument that high volatility periods are "buying opportunities" for long horizon investors. It also enables our stochastic model for returns to fit the most salient features of the US data, i.e. that both expected returns and volatility vary significantly over time (Campbell and Shiller, 1988; Schwert, 1989) but are not strongly related to each other at short horizons, despite the fact that increases in volatility are associated with market downturns (Glosten et al., 1993). Finally, it allows us to compare the utility benefits from timing variation in volatility to the long literature on the utility benefits of timing expected returns (for example, Campbell and Viceira (1999), Barberis (2000)). While the benefits from expected return timing have been studied extensively, the potential benefits from volatility timing have received much less attention. In our analysis, we also use both the parameter point estimates and the associated estimation uncertainty to consider a range of parameters governing the return process that are likely given the data.

Given this return process, we study the portfolio problem of an infinite-lived investor with recursive preferences (Epstein and Zin, 1989) with unit EIS. These preferences allow us to conveniently control the horizon of the investor, i.e. the timing of her consumption, while at the same time it also keeps the environment stationary and not time dependent. These preferences should accurately capture individuals and institutions that target a constant expenditure share of their wealth (e.g., university endowments, sovereign wealth funds, or pension funds). Given investor preferences and the return process, we then quantitatively study how the optimal portfolio responds to volatility.

As is typical in the portfolio choice literature (Merton, 1971) our optimal portfolio weight in the risky asset takes the form

$$(portfolio\ weight)_t = (myopic\ demand)_t + (hedging\ demand)_t, \tag{1}$$

where the myopic demand, $\mu_t / \gamma \sigma_t^2$, is equal to the optimal portfolio weight of a short horizon, log utility, or mean-variance investor.

In light of this equation, our quantitative questions are (1) $\partial(portfolio\ weight)/\partial\sigma^2$:

what is the optimal response to a change in volatility?, and (2) what are the utility costs associated with ignoring variation in volatility (i.e., what are the costs of a portfolio strategy that sets $\partial(portfolio\ weight)/\partial\sigma^2 = 0$)?

The effect of volatility on the myopic, mean-variance demand term is strongly negative: in the data increases in variance are not offset by proportional increases in expected returns, so an increase in variance lowers the myopic demand. Our estimation finds that expected returns do rise by small amounts after an increase in volatility, but this is not nearly enough to keep the term μ_t/σ_t^2 from falling; that is, the elasticity of this term with respect to volatility is near -1. Moreira and Muir (2016) empirically show that volatility timing can increase Sharpe ratios through this channel for a wide range of factors. Thus, short horizon investors or investors with log utility, for which the hedging demand term is absent, should sell in response to an increase in volatility.

The second term, the hedging demand term, relates primarily to the amount of mean-reversion in stock returns and can be quantitatively large on average, particularly for longer investment horizons, when risk aversion is greater than 1 – a result which has been extensively analyzed in the literature (e.g., Campbell and Viceira (1999), Brandt (1999) Barberis (2000) and Wachter (2002)).[1] However, there is little work on how this term may change in response to changes in volatility. Thus, rather than focusing on the level of the hedging demand, we are interested in its dynamics and in particular in how the hedging demand term changes with volatility: $\partial(hedging\ demand)_t/\partial\sigma_t^2$.

In fact, there is a widespread consensus among practitioners and academics that variation in the hedging demand term is such that long-term oriented investors should not volatility time at all. For example, Cochrane (2008a), Buffett (2008), and more recently Vanguard –a leading mutual fund company– argue that long-term oriented investors are better off ignoring movements in volatility.[2] The argument is that, since volatility is typically associated with market downturns, and downturns are attractive buying opportunities, it is not wise to sell when volatility spikes. Further, and more importantly, because of

[1] There is also a hedging demand term that relates to volatility shocks, but quantitatively this turns out to be small (Chacko and Viceira, 2005).

[2] For the Vanguard reference see "What to do during market volatility? Perhaps nothing." See `https://personal.vanguard.com/us/insights/article/market-volatility-082015`. In the appendix, we reference similar advice from Fidelity, the New York Times, and many other sources.

mean-reversion in stock returns, investors with long horizons should not view increases in volatility as an increase in risk – the idea is that an increase in volatility makes stock prices more uncertain tomorrow, but not more uncertain over long horizons that these investors care about. Thus, the argument is that periods of high volatility may be much more attractive to long horizon investors relative to short horizon investors through the hedging demand term.

We find that if the share of mean-reverting shocks is constant (i.e., the volatility of mean-reverting shocks increases proportionally to total return volatility), as it is typically assumed in the literature, then the hedging demand is essentially constant, meaning it does not change with volatility. Specifically, we show that a simple strategy of the form $w^* = \omega_0 + \frac{1}{\gamma}\frac{\mu_t}{\sigma_t^2}$ achieves the same utility as the fully optimal strategy, so that the optimal portfolio can be approximated very accurately by the myopic portfolio plus a constant weight investment in the buy-and-hold portfolio. Thus, contrary to conventional wisdom, investors with different horizons should reduce their dollar investment in equity by exactly the same amount in response to changes in the risk-return trade-off. This affine form for the portfolio strategy holds for a wide range of parameters that are likely given the data. Our answer to question (1) is thus unambiguous. A long-lived investor should volatility time quite aggressively. It is worth emphasizing that, though our rich process for return dynamics requires a numerical solution for the optimal portfolio, the optimal portfolio is almost perfectly approximated with a simple, practical, and intuitive portfolio rule. This differs from Campbell and Viceira (1999) who first approximate the portfolio problem itself through log-linearization and then study an analytical solution.

We next evaluate the utility benefits from volatility timing, where we define a volatility timing strategy as a strategy that *only* uses conditional information on volatility, but not expected returns. Specifically, we restrict ourselves to constant weight combinations of the buy-and-hold portfolio and the volatility managed portfolio from Moreira and Muir (2016), i.e. strategies of the form $w^\sigma = \omega_0 + \omega_1\frac{1}{\gamma}\frac{\overline{\mu}}{\sigma_t^2}$ where $\overline{\mu}$ sets the expected return to its unconditional mean. Notice, this strategy is exactly the fully optimal strategy described before but does not time conditional expected returns. We compare the utility of this strategy to the fully optimal strategy, w^*, that conditions on both expected returns and volatility, and to the optimal buy-and-hold strategy that chooses a constant weight in the risky-asset, $\overline{w}$.

We find very large gains from volatility timing. The increase in utility from volatility timing relative to a buy-and-hold strategy ranges from 20% to 80% with our point estimate implying gains close to 35%. These gains are about 60% of the *total* gain of switching from the buy-and-hold strategy $\overline{w}$ to the fully optimal strategy w^* (i.e., a strategy that also conditions on expected returns as well as volatility). Thus, ignoring variation in volatility is very costly, even for long horizon investors, and the benefits to timing volatility are significantly larger than the benefits to timing expected returns. We then reevaluate these gains taking into account the parameter uncertainty implied by our estimation procedure for return dynamics. Specifically, we use the uncertainty in our estimation to recover the probability that inaction is the optimal response to volatility variation and find that this probability is close to zero. We then use the estimation uncertainty to evaluate the robustness of the gains from volatility timing. We find that the the gains vary as function of the parameters but are extremely likely to be positive and large. In contrast, we find that the gains from expected return timing are much more sensitive to parameter uncertainty, consistent with Barberis (2000) and Pástor and Stambaugh (2012) among others.

We then relax the standard assumption that the share of mean-reverting shocks is constant and show that variation in the composition of volatility leads to variation in the hedging demand term. This volatility composition channel could act as a counteracting force to the variation in the myopic demand. This happens when increases in volatility are associated with a larger share of mean-reverting shocks; that is, prices become more volatile only in the short term but not more volatile in the long term because of an increased degree of mean-reversion. Both Cochrane (2008a) and Buffett (2008) argued that the huge spike in volatility in the fall of 2008 was mostly about "short-term volatility".[3] Motivated by this, we complement our analysis by allowing the composition of volatility shocks to be time-varying. In particular, allowing for a positive correlation between

[3] "And what about volatility?(...) expected returns would need to rise from 7% per year to 78% per year to justify a 50/50 allocation with 50% volatility. (...) The answer to this paradox is that the standard formula is wrong. (...) Stocks act a lot like long-term bonds – (...)If bond prices go down more, bond yields and long-run returns will rise just enough that you face no long-run risk.(...)the same logic explains why you can ignore "short-run" volatility in stock markets."(Cochrane, 2008a)

volatility and the share of mean-reverting shocks allows us to study the notion that investors should ignore "short-term volatility."

To explore this case, we set the volatility of permanent shocks to returns (sometimes labeled "cash flow shocks") to be constant. We then have all time-variation in return volatility be driven by the volatility of transitory or mean-reverting shocks (labeled "discount rate shocks"). Thus, when volatility is very low, returns are entirely driven by permanent shocks, i.e., there is no mean-reversion in returns and no return predictability. Then both long term and short term investors will choose the same allocation to stocks (Samuelson, 1969) – the hedging demand term will be zero. However, in high volatility times, stock returns become strongly mean-reverting because the volatility of discount rate shocks increases. In these periods, a short term investor sees the increase in volatility and wants to sell. The long term investor weighs two effects: the myopic desire to sell, but also the large hedging demand that now arises from mean-reversion. Thus, the long term investor will react less strongly to the increase in volatility in this case, because it is accompanied by an increase in the degree of mean-reversion in returns. Here, hedging demands are no longer constant, but are positively correlated with volatility.

Empirically, there is no evidence on how the share of mean-reverting shocks varies with volatility. Importantly, even in the case of extreme co-movement described above, when volatility variation is completely driven by variation in the volatility of mean-reverting shocks, we show that long-term investors still find it optimal to time volatility. The key for this result is that in the data mean reversion takes many years, making even mean-reverting shocks risky for realistic investment horizons.[4] Specifically, we find that now the optimal portfolio has a weight on the volatility timing portfolio that is 30% lower than before, meaning investors time volatility somewhat less aggressively, and we find that volatility timing can capture utility gains of around 20% relative to a buy-and-hold strategy.

Our results are important for investors such as pension funds, endowments, sovereign wealth funds, individuals saving for retirement, or other long-term investors as they pro-

[4]Sharpe ratios for stocks increase only slowly with investment horizon (Poterba and Summers, 1988), and valuation ratios that predict returns are highly persistent with auto-correlation close to one (Campbell and Shiller, 1988).

vide guidance in how to optimally respond to volatility. To the best of our knowledge this paper is the first to directly speak to the consensus view that long term investors should ignore volatility variation, and is the first to study the portfolio choice response to volatility in the presence of expected return shocks. We also highlight how variation in the composition of volatility shocks could potentially make long-horizon investors optimally ignore variation in volatility. Finally, we argue that our results hold up for other preference specifications and are likely to hold up in settings with non-financial income. Our results are also important because they further sharpen the puzzle documented in Moreira and Muir (2016). That paper finds that short-term investors should sell when volatility increases. The results in this paper show that a longer investment horizon does not qualitatively change the desire to sell after an increase in volatility. Thus, horizon effects cannot explain the weak equilibrium relationship between expected returns and volatility Moreira and Muir (2016) document.

The paper proceeds as follows. Section 2 describes the process for returns and investor preferences. Section 3 analyzes the optimal portfolio and associated utility gains from volatility timing. Section 4 describes our parameter estimation in more detail and studies robustness of our results to parameter uncertainty. Section 5 contains extensions to our main results. Section 6 concludes.

1. Literature Review

Our paper builds on the prolific literature on long-term asset allocation. Starting with the seminal work of Samuelson (1969) and Merton (1971), this literature has studied carefully the implications of mean-reversion for portfolio choice. Campbell and Viceira (1999), Barberis (2000) and Wachter (2002) study the optimal portfolio problem in the presence of time-varying expected returns. The key result is that the presence of mean-reversion in market returns imply investors with longer horizons should invest more in the stock market. An important caveat is that parameter uncertainty can attenuate these horizon effects (see Barberis (2000) and Xia (2001)). To a large extent, the results in this literature have percolated into practice and non-academic discourse. The view that the stock market is safer over the long-term is now standard in the money management industry. Also in line with this literature is the view that market dips are good buying opportunities (Campbell and Viceira, 1999).

Much less studied, but we think equally important, is time-variation in second moments. Chacko and Viceira (2005) and Liu (2007) study variation in volatility and Buraschi et al. (2010) study variation in correlations. For realistic calibrations, this literature finds only modest deviations from myopic behavior. Thus, the optimal portfolio is very close to the simple myopic weight. The absence of large hedging demands suggests that volatility timing as in Moreira and Muir (2016) is desirable and investment horizon effects are not first order. However, these papers abstract from variation in expected returns. Thus, they cannot speak to the conventional wisdom that volatility spikes are mostly "buying opportunities" or that return volatility is mostly due to transitory shocks that mean-revert over the long run. It is precisely this gap that this paper fills. Consistent with the intuition behind the traditional view, we show that there is an important interaction between volatility and expected return variation through the volatility composition channel. However, we show that for parameters consistent with the data, this mechanism is not large enough to offset the variation in the myopic demand. Related papers that account for both volatility and expected returns include Collin-Dufresne and Lochstoer (2016) and Johannes et al. (2014). Collin-Dufresne and Lochstoer (2016) have a time-varying risk-return relationship in a general equilibrium setting and point out that long-terms investors only want to buy at "low prices" if effective risk-aversion, rather than risk itself, has increased in order to cause the fall in prices. Johannes et al. (2014) solve a Bayesian problem that accounts for time-varying volatility when forming out of sample expected return forecasts.

Finally, we build on the results in Moreira and Muir (2016) who study volatility timing in the context of a mean-variance investor who simply maximizes unconditional Sharpe ratios. This paper generalizes those results by allowing for much more general preferences, and goes well beyond the results in that paper by solving for the optimal portfolio, considering parameter uncertainty, and by studying the interaction of volatility and mean-reversion in returns.

2. The portfolio problem

We study the problem of a long-horizon investor and investigate how much they should adjust their portfolio to changes in volatility.

2.1 Investment opportunity set

We assume there is a riskless bond that pays a constant interest rate r, and a risky asset S_t, with dynamics given by

$$\frac{dS_t}{S_t} = (r + \mu_t)dt + \sigma_t dB_t^S, \tag{2}$$

where S_t is the value of a portfolio fully invested in the asset and that reinvests all dividends. We model expected (excess) returns as an auto-regressive process with stochastic volatility,

$$d\mu_t \quad = \quad \kappa_\mu(\overline{\mu} - \mu_t)dt + \sigma_\mu \sigma_t dB_t^\mu, \tag{3}$$

Notice that this means that the volatility of shocks to expected returns scale up and down proportionally with shocks to realized returns. In later sections, we consider cases where we break this proportionality. We write log volatility $f(\sigma_t^2) = \ln\left(\sigma_t^2 - \underline{\sigma}^2\right)$ as an auto-regressive process with constant volatility,

$$df(\sigma_t^2) \quad = \quad \kappa_\sigma \left(\overline{f} - f(\sigma_t^2)\right) dt + \nu_\sigma dB_t^\sigma, \tag{4}$$

where the parameter $\underline{\sigma}^2$ controls the lower bound of the volatility process. This lower bound is important in eliminating arbitrage opportunities (i.e., infinite Sharpe ratios). Our assumption about a lognormal volatility process should not be seen as crucial, although it allows for easier solutions in our numerical exercise. Results using a square root process (Heston, 1993; Cox et al., 1985) for volatility along the lines of Chacko and Viceira (2005) are similar.

Shocks to realized returns, expected returns, and volatility, are thus captured by the

Brownian motions dB_t^S, dB_t^μ, and dB_t^σ. We now specify the correlation of these shocks. First, we impose

$$E_t[dB_t^\mu dB_t^S] = -\frac{\sigma_\mu}{\kappa_\mu}, \tag{5}$$

Note that the correlation between between expected returns and realized returns is not a free parameter. The correlation $-\frac{\sigma_\mu}{\kappa_\mu}$ implies that shocks to expected returns induce an immediate change in prices so that in the long run, it exactly offsets expected return innovations, i.e. it imposes that expected return shocks have no effect on the long-run value of the asset. We make this choice to emphasize that we want to consider transitory shocks to returns that have no long run impact, however, we also note that if one freely estimates this correlation in the data, one recovers roughly this value (Cochrane, 2008b) – hence it is not an overly restrictive assumption. This correlation also defines the share of "discount rate shocks" that drive returns – that is, when the correlation is 1, then all variation in returns is driven by discount rate shocks, and when it is zero, expected return shocks play no role. We label this correlation $-\alpha_\mu$. We will thus write $\sigma_\mu = \alpha_\mu \kappa_\mu$ and focus on estimating α_μ and κ_μ in the data as these parameters have direct economic interpretations as the share and persistence of discount rate shocks.

We next specify parameters the remaining correlations

$$E_t[dB_t^\mu dB_t^\sigma] = \rho_{\sigma,\mu}, \tag{6}$$

$$E_t[dB_t^S dB_t^\sigma] = -\rho_{\sigma,\mu}\alpha_\mu - \rho_{\sigma,S}\sqrt{1-\alpha_\mu^2}, \tag{7}$$

The correlation between volatility and expected and realized returns are free parameters which must satisfy $\rho_{\sigma,S}^2 + \rho_{\sigma,\mu}^2 \leq 1$.[5] Finally, we set the unconditional mean of the log volatility process $f(\sigma_t^2)$ to $\overline{f} = \ln(\overline{\sigma}^2 - \underline{\sigma}^2) - \frac{\nu_\sigma^2}{2\kappa_\sigma}$.

This parametrization leads to a natural interpretation of the parameters: $\overline{\mu}$ is the average expected excess return of the risky asset, $\overline{\sigma}^2$ is the average conditional variance of

[5] We specify these correlations as constant. In particular, we don't consider time-variation in the correlation between volatility and expected returns. See Collin-Dufresne and Lochstoer (2016) for a case where this time-variation plays a role in a general equilibrium model for long term portfolio choice.

returns, ν_σ^2 is the conditional variance of log variance, $\rho_{\sigma,\mu}$ controls the covariance variance and discount rate shocks, $\rho_{\sigma,S}$ controls the covariance between variance and cash flow shocks (return innovations uncorrelated to innovation is discount rates). Throughout, we adopt the language from the literature (Campbell and Shiller, 1988; Campbell, 1996; Campbell and Vuolteenaho, 2004), using "cash flow shocks" to denote permanent shocks to returns that are uncorrelated to shocks that affect expected returns. Next, α_μ denotes the discount rate share of return variation.

The stochastic environment described by (2) and (3) allows for variation in volatility; variation expected returns (i.e., mean-reversion in returns); and flexible time-series relation between expected returns and volatility ($\rho_{\sigma,\mu}$). The latter governs the risk-return trade-off relationship between variance and the risk premium. In the appendix, we discuss even more sophisticated and flexible ways of modeling this relationship. In particular, we discuss allowing expected returns to more directly depend on volatility by having two frequencies for expected returns: a shorter frequency component that is related to volatility, and a slower moving component (specifically, we write $\mu_t = x_t + b\sigma_t^2$ where b governs the risk-return relation and x and σ^2 are allowed to move at different frequencies). It turns out, however, that because the risk-return relation is empirically weak, we do not lose much by incorporating a less rich relationship between expected returns and variance. In fact, we will show in our estimation that the current model is able to capture the essential empirical moments relating risk and return in the time-series, meaning our modeling of the risk-return tradeoff is appropriate. Finally, in later we also allow for variation in the composition of volatility shocks (that is, we consider the case where α_μ is not constant, but varies over time). This will allow for variation in the share of return volatility due to discount-rate shocks.

Together, these ingredients are novel and essential to study the optimal response to volatility variation. Earlier work on portfolio choice has studied expected return variation, volatility variation, or volatility variation with a constant risk-return trade-off. Examples of work that study volatility timing in a dynamic environment are Chacko and Viceira (2005) and Liu (2007). But these papers do not study the interaction of discount rate and volatility shocks which are the basis for the conventional view that long horizon investors should ignore volatility variation.

2.2 Estimation of parameters

We estimate the model using Simulated Method of Moments (Duffie and Singleton, 1993) and use the estimated parameters in Table 1 to discuss the model implications for portfolio choice.

Our goal is for the model to match the key dynamic properties of US stock returns documented in the empirical finance literature. With that in mind, we use the US market excess return from 1926-2015 from Ken French (ultimately, the CRSP value-weighted portfolio). We use daily data to construct a monthly series of realized volatility, RV, which will use to match the properties of volatility in the model. Specifically, we will simulate the model at daily frequency and compute realized volatility in the same manner as in the data – thus the true volatility process is unobserved. We then aggregate to monthly frequency in the data and model to match all moments. Finally, we bring in additional monthly data on the US dividend price ratio from Robert Shiller to match moments related to expected returns and return predictability.

We first calibrate the real riskless rate ($r = 1\%$) and the market expected excess returns ($\overline{\mu} = 5\%$), which reflect the U.S. experience in the post-war sample. We also calibrate the volatility lower bound to ($\underline{\sigma} = 7\%$) based on the data.[6]

We estimate the remaining seven parameters. Let $\theta = (\overline{\sigma}^2, \nu_\sigma, \kappa_\sigma, \alpha_\mu, \kappa_\mu, \rho_{\sigma,\mu}, \rho_{\sigma,S})$ be the vector of parameters to be estimated. Our SMM estimator is given by

$$\hat{\theta} = \arg\min_{\theta}(g(\theta) - g_T)' S(\theta)(g(\theta) - g_T), \tag{8}$$

where g_T is a set of target moments in the data and $g(\theta)$ is the vector of moments in the model for parameters θ. We use an identity weighting matrix S in our main results.

We choose the vector of target moments g_T to be informative about the parameters θ. Our target moments are: (1) average realized monthly variance, (2) the auto-correlation

[6] Here we use that the minimum of the VIX from 1990-2015 is 10%, so our 7% for the longer 90 year sample is reasonable. Note we use VIX to calibrate this number rather than realized volatility, because realized volatility is noisy and hence would not properly measure a lower bound for true volatility.

coefficient of the logarithm of realized monthly variance, (3) the standard deviation of innovations to log realized variance based on an AR(1) forecasting model, (4) the covariance between volatility innovations and realized returns, (5) the alpha of the volatility managed market portfolio on the market portfolio (see (Moreira and Muir, 2016)), (6-7) the R-squared of a predictability regression of one month and five year-ahead returns on the price-dividend ratio. The the alpha of the volatility managed portfolio is defined by the regression $\frac{c}{RV_t^2}R_{t+1} = \alpha + \beta R_{t+1} + \varepsilon_{t+1}$ where the alpha measures whether one can increase Sharpe ratios through volatility timing. Moreira and Muir (2016) show this alpha measures the strength of the risk-return tradeoff over time, but is a sharper measure than standard forecasting regressions.

While there is not an exact one-to-one mapping between moments and parameters, the link between parameters and moments is intuitive, and the moments are very informative about the parameters of interest. Average realized monthly variance identifies $\overline{\sigma}^2$. The auto-correlation of volatility and the standard deviation of volatility innovations identify ν_σ and κ_σ. These moments imply that the estimated volatility process is highly volatile but not very persistent. The return predictability R-squares at one month and five year horizons identify α_μ, the discount rate share, and κ_μ, the volatility and persistence of discount-rate shocks. Intuitively, the one-month R-square implies the share of discount-rate shocks must be large and the fact that five-year R-squares are substantially larger implies that expected returns must be highly persistent. The covariance between realized returns and volatility innovations and the volatility managed alpha identify $\rho_{\sigma,\mu}$ and $\rho_{\sigma,S}$. In the data, the large negative correlation between volatility innovations and realized returns implies that $\rho_{\sigma,\mu} + \rho_{\sigma,S}$ is close to one. The alpha of the volatility managed portfolio disciplines the extent to which this co-movement is due to a correlation between discount rates and volatility shocks. In the data, a portfolio that takes less risk when volatility is high generates a large Sharpe ratio, implying that the co-movement between volatility and discount rate shocks is not strong (see Moreira and Muir (2016)).

Table 1 reports targeted moments in the model and in the data. Overall the model matches the data extremely well and matches the key empirical facts on the dynamics of stock returns documented in the finance literature. In particular, the estimated volatility process is highly volatile, so there substantial time-variation in conditional volatility (Schwert, 1989). Expected returns are quite variable, i.e. discount-rate volatility is an important component of stock market volatility (Campbell and Shiller, 1988), and these discount rate shocks are strongly correlated with volatility shocks (French et al., 1987). That is, increases in volatility are associated with low realized returns and increases in expected returns. However, this correlation does little to dampen variation in the risk-return trade-off because shocks to expected returns are much more persistent than shocks to volatility, and also the correlation between volatility and expected returns, while large, is not equal to 1. Thus, the model is able to produce positive volatility managed alphas consistent with Moreira and Muir (2016) because the model, like the data, does not feature an overly strong risk return tradeoff.[7] That is, consistent with a long literature, there is some risk-return tradeoff in the data but it appears to be fairly weak (French et al., 1987; Glosten et al., 1993; Lettau and Ludvigson, 2003). Thus, taken together, our process for returns matches the key empirical features about the properties of expected returns, conditional volatility, and realized returns documented by a long literature in asset pricing.

We also report bootstrapped standard errors for the estimated parameters in Table 1. That is, we reestimate the model using many 90 year simulations and reestimate parameters to have a sense of parameter variation. We report standard deviations across parameter estimates obtained from moment matching individual simulations. Consistent with the large literature on market timing, the dynamics of expected returns is the least well estimated aspect of our model. This estimation uncertainty will play a role in later sections where we consider that the investor may not know the true parameters in making his portfolio decision.

2.3 Preferences and optimization problem

Investors preferences are described by Epstein and Zin (1989) utility, a generalization of the more standard CRRA preferences that separates risk aversion from elasticity of in-

[7] We undershoot slightly the volatility managed alpha because we calibrate the equity premium to 5%, which is lower than the in sample equity premium (7.8%). In unreported results we verify that our model generates a 5% volatility managed alpha if we were to calibrate the model to the in sample equity premium.

tertemporal substitution. We adopt the Duffie and Epstein (1992) continuous time implementation and focus on the case of constant elasticity of substitution:

$$J_t = E_t \left[\int_t^{\infty} f(C_s, J_s) ds \right], \tag{9}$$

where $f(C_t, J_t)$ is an aggregator of current consumption and continuation utility that takes the form

$$f(C, J) = h(1-\gamma)J \times \left[log(C) - \frac{log((1-\gamma)J)}{1-\gamma} \right], \tag{10}$$

where h is rate of time preference, γ the coefficient of relative risk aversion. The unit elasticity of substitution is convenient for our purposes because it allow us to directly vary the investor horizon in a way that is independent of the attractiveness of the investment opportunity set. Specifically, $1 - \exp(-h)$ is the share of investors wealth consumed within one year. Thus $1/h$ can be thought as the horizon of the investor. In Section 5.1 we consider alternative preference specifications.

The investor maximizes utility subject to his intertemporal budget constraint (Eq. 11 below) and the evolution of state variables (Eq. (3)). Let W_t denote the investor wealth and w_t the allocation to the risky asset, then the budget constraint can be written as,

$$\frac{dW_t}{W_t} = w_t \left(\frac{dS_t}{S_t} - rdt \right) + rdt - \frac{C_t}{W_t} dt. \tag{11}$$

3. Analysis

Our aim is to quantify the optimal amount of volatility timing for a realistic portfolio problem in which an investor decides how much to invest in the market portfolio and in a riskless asset. We solve for the investor value function numerically and study how the optimal portfolio should respond to changes in volatility. Our analysis is quantitative in nature and it is therefore important that our model for returns described in Eqs. (2)-(3) fit the dynamics of returns in the data.

In the baseline case we study the problem of an investor with a 20 year horizon ($h =$

1/20) and risk-aversion of 5, and we investigate the sensitivity of our results to these parameter choices.

3.1 Solution

The optimization problem has three state variables: the investor's wealth plus the investment opportunity set state variables μ_t, σ_t.

The Bellman equation for this problem is standard

$$
\begin{aligned}
0 &= \sup_{w,C} \{ f(C_t, J_t) + [w_t \mu_t W_t + r W_t - C_t] J_W + \frac{1}{2} w_t^2 W_t^2 J_{WW} \sigma_t^2 && (12)\\
&- w_t W_t \left(J_{W\mu} \alpha_\mu \sigma_t^2 + J_{W\sigma} \sigma_t \left(\alpha_\mu \rho_{\sigma,\mu} + \sqrt{1-\alpha_\mu^2} \rho_{\sigma,S} \right) \right) && (13)\\
&+ J_\mu \kappa_\mu (\overline{\mu} - \mu_t) + J_\sigma \kappa_\sigma (\overline{f} - f(\sigma_t^2)) + \frac{1}{2} \left(J_{\mu\mu} \alpha_\mu^2 \kappa_\mu^2 \sigma_t^2 + J_{\sigma\sigma} \nu_\sigma^2 + 2 J_{\mu\sigma} \rho_{\sigma,\mu} \nu_\sigma \sigma_t \alpha_\mu \kappa_\mu \right) \},
\end{aligned}
$$

where we omit the argument on $J_t = J(W_t, \mu_t, \sigma_t)$ for convenience. It is well known that the value function for this type of problem is of the form $J(W, Z) = \frac{W^{1-\gamma}}{1-\gamma} e^{g(\mu_t, \sigma_t)}$. Plugging this form in (12) we obtain that the optimal consumption to wealth ratio is constant, $C_t = hW_t$ and the optimal portfolio weight satisfies

$$w^*(\mu_t, \sigma_t^2) = w^m(\mu_t, \sigma_t^2) + w^h(\mu_t, \sigma_t^2), \tag{14}$$

where the first term in (14) is the myopic portfolio weight

$$w^m(\mu_t, \sigma_t^2) = \frac{1}{\gamma} \frac{\mu_t}{\sigma_t^2}. \tag{15}$$

It calls the investor to scale up his position on the risky asset according to the strength of the risk-return trade-off and her coefficient of relative risk aversion. This is also the optimal portfolio weight of a short-horizon mean-variance investor (or log-investor). The additional term in (14) is a hedging demand (Merton, 1971) which is given by

$$w^h(\mu_t, \sigma_t^2) = -\frac{1}{\gamma} g_\mu \alpha_\mu - \frac{1}{\gamma} g_\sigma \frac{\left(\alpha_\mu \rho_{\sigma,\mu} + \sqrt{1-\alpha_\mu^2} \rho_{\sigma,S} \right)}{\sigma_t}. \tag{16}$$

The hedging demand w^h arises because a long-horizon investor is concerned with the overall distribution of her consumption and not only the short-term dynamics of her wealth. Changes in the risky asset expected returns or volatility lead to changes in the distribution of the investor wealth, resulting in a demand for assets that hedge these changes. To the extent that the risky asset is correlated with changes in the opportunity set , this demand for hedging impacts the investor's position in the risky asset.

This hedging effect means that a long-horizon investor might behave very differently from a short-term oriented investor. An increase in volatility might generate an increase in the hedging demand that is enough to completely offset the reduction in exposure due to the myopic demand– i.e. it might be that long-horizon investors should just ignore time-variation in volatility, in line with the argument articulated in Cochrane (2008a).[8]

The empirical fact that expected returns increase after low return realizations, $dB^{\mu}dB^{s} < 0$, makes investment in the risky asset a natural investment hedge for changes in expected returns. This effect has been studied extensively in the literature (e.g. Campbell and Viceira (1999), Barberis (2000), and Wachter (2002)), which has shown that when $\gamma > 1$, this hedging demand leads a long-horizon investor to have a larger average position in the risky asset.

A similar hedging demand arises due to changes in volatility, though with the opposite sign. The fact that increases in volatility tend to be associated with low return realizations also implies that the risky asset co-moves with the investment opportunity set. Work by Chacko and Viceira (2005) and more recently Buraschi et al. (2010) show that this effect tends to be small for realistic calibrations, which we confirm here for realistic parameters. Specifically, when $\gamma > 1$ the hedging demand due to volatility pushes investors to hold slightly smaller positions in the risky asset.

The direction of these hedging demands follows from the interaction between changes in the Sharpe ratio and the coefficient of relative risk-aversion. An investor that is more

[8]"And what about volatility?(...) if you were happy with a 50/50 portfolio with an expected return of 7% and 15% volatility , 50% volatility means you should hold only 4.5% of your portfolio in stocks! (...) expected returns would need to rise from 7% per year to 78% per year to justify a 50/50 allocation with 50% volatility. (...) The answer to this paradox is that the standard formula is wrong. (...) Stocks act a lot like long-term bonds – (...)If bond prices go down more, bond yields and long-run returns will rise just enough that you face no long-run risk.(...)the same logic explains why you can ignore "short-run" volatility in stock markets."(Cochrane, 2008a)

conservative than a log investor ($\gamma > 1$) wants to transfer resources from states where the opportunity set is better to states where it is worse. Because expected and realized returns are negatively correlated, a positive tilt towards the risky asset implies her wealth increases following reduction in the Sharpe ratio due to a reduction in expected returns. Symmetrically, because volatility and realized returns are negatively related, a *negative* tilt towards the risky asset implies her wealth increases following a reduction in the Sharpe ratio due to an increase in volatility.

Investment horizon, together with the persistence of the state variables ($\kappa_\mu, \kappa_\sigma$), shapes the strength of the hedging demand through the sensitivity of the value function to changes in the state variables (g_μ, g_σ). Intuitively, persistent changes to the state impact the investment opportunity set for longer, and this impact is larger for investors with a longer horizon, which are naturally more exposed to persistent changes in the opportunity set. As a result the value function is typically more sensitive to the state variables and the resulting hedging demands are larger for investors with longer horizons. Here the unit IES is particularly convenient as the patience parameter h directly controls the effective horizon of the investor, i.e. the timing of their consumption.

We use projection methods to solve for $g(\mu_t, \sigma_t)$. See Appendix for details.

3.2 Optimal portfolios

It is illuminating to discuss our results by contrasting the optimal choices of long and short-term investors. Because we are especially interested in how investors should respond to variation in volatility, we first represent our results in terms of an Impulse Response Function (IRF). In the top panels of Figure 1 we start by showing the response of variance and expected returns to a one standard deviation shock to variance, and then show how long and short-term investors respond.

Expected returns go up in response to a volatility shock, though quantitatively this increase is small. This is due to the high correlation between realized returns and volatility innovations present in the data. Thus, a innovation in volatility is correlated with innovations in expected returns. Nevertheless, the myopic and the optimal portfolio go down sharply and in parallel. This means that two investors with the same risk-aversion but different horizons will reduce the fraction of their wealth allocated to stocks by exactly the same amount. In this sense, horizon has no impact on how investors should respond to changes in volatility.

There are however large level differences across portfolios. The long term investors invests on average a much higher fraction of their wealth in stocks. Level differences across the optimal and the myopic portfolio are shown in the level of the flat yellow line, which plots $w^h(\mu_t, \sigma_t^2) = w^*(\mu_t, \sigma_t^2) - w^m(\mu_t, \sigma_t^2)$ normalized by the steady state long-term portfolio $w^*(\overline{\mu}, \overline{\sigma}^2)$. The long-term investor has a risk exposure that is about 20% larger than the myopic investors in the steady state, but this difference–as fraction of the risky portfolio share– grows large as volatility goes up and the myopic weight goes down.

The flat yellow line implies that the hedging demand is, at least locally, not related to volatility. The hedging demand term drives a difference between short and long term investors, but this hedging demand term is roughly constant, so that conditional responses to volatility variation are not substantially different. The response of a long-term investor to volatility is completely driven by the myopic component of her portfolio, i.e. variation due to the instantaneous risk-return trade-off.

3.2.1 The optimal portfolio is simple

Motivated by the constant hedging term we see in Figure 1, we consider portfolio strategies that invest in the myopic portfolio plus a constant position in the buy-and-hold portfolio,

$$\widetilde{w}^*(\mu_t, \sigma_t^2) = \omega_0 + \omega_1 w^m(\mu_t, \sigma_t^2). \tag{17}$$

We solve for the investor's lifetime utility and find that a portfolio strategy with $\omega_0 = E[w_t^h]$ and $\omega_1 = 1$ attains the same life-time utility as the optimal portfolio. The approximation $\widetilde{w}^*$ is not only a good local approximation for the optimal portfolio, but also an excellent global approximation. We refer to $\widetilde{w}^*(\mu_t, \sigma_t^2)$ as the optimal linear portfolio because it's weight is a linear function of the myopic portfolio.[9]

This result can be seen in Table 2 which shows the optimal policy weights and the

[9]Formally, it is an Affine function of w^m.

percentage lifetime expected utility loss from switching from the optimal portfolio to the Affine approximation. A utility loss close to zero implies the investors forego almost no consumption if it adopts the simpler strategy. Thus $\widetilde{w}^*$ provides a good global approximation to w^*.[10]

The results that the optimal linear portfolio $\widetilde{w}^*$ achieves the optimal utility is important because the numerical solution generates simple and implementable portfolio advice. Every Investor can implement their strategy with two mutual funds–one that holds the market and one that times the risk-return trade-off. The results that $\omega_1 = 1$ implies that investment horizons play a role only on the allocation to the buy-and-hold mutual fund. At least for our point estimates, all investors, irrespective of their investment horizon, allocate the same fraction of their wealth to the timing mutual fund.

Table 2 also shows that the (approximate) optimality of the linear portfolio holds up across a wide range of parameters for the stochastic process, investment horizons, and risk aversion. Thus, investor portfolio response to volatility–as a fraction of their wealth–is always the same regardless of the investment horizon.

3.2.2 The optimal portfolio elasticity to changes in volatility

Another way to evaluate how responsive to volatility changes investors should be is in terms of an elasticity, i.e. the percentage change in the portfolio allocation resulting from a 1% increase in volatility,which is defined as

$$\zeta = -\frac{dlog(w^*(\mu_t, \sigma_t^2))}{dlog(\sigma_t^2)}. \tag{18}$$

This perhaps provides a more direct measure of the importance of volatility driven changes for a particular investor. For a myopic investor $\zeta = 1 - d\ln(\mu_t)/d\ln(\sigma_t^2)$, which goes to 1 as the conditional risk-return trade-off goes to zero. Our estimates imply $\zeta^m = 0.97$, which reflects the small increase in expected return following a volatility shock we see in Figure 1. An elasticity of 1 implies an investor reduce their exposure to stocks by 10% for a 10% increase in volatility.

[10] A 1% utility loss is equivalent to decreasing the investor consumption by 1% state by state.

The approximation (17) implies $\zeta \approx \frac{w_t^m}{\omega_0 + w_t^m}\zeta^m$, with the long-term investor elasticity lower than the myopic as long $\omega_0 > 0$. The elasticity also goes to zero as the myopic weight goes down, due for example to an increase in volatility. In Table 2 we focus on the elasticity of the optimal portfolio around the median value of the state variables to a one-standard deviation increase in variance,i.e. the typical response to volatility. For the baseline parameters we find an elasticity of 0.7, which implies that as a share of her portfolio a long horizon investor should respond less aggressively to variation in volatility. This happens because the long-horizon investor has a larger investment in the stock market to begin with (from the hedging demand term).

In Table 2 we see that variation in ζ tracks variation in ω_0, the optimal allocation to the buy-and-hold portfolio. For example, when the expected return is very volatile , high α_μ, the weight ω_0 is extremely high and the elasticity very low, close to 0.4. To a smaller extent this also happens as we increase the investment horizon with the elasticity going down from 0.68 to 0.65 as the investment horizon increases from 20 to 50 years.

In summary, the data on stock market returns when looked at through the lens of the standard moments studied in the literature, strongly rejects the conjecture that investors should ignore movements in volatility. Investors with long investment horizons are somewhat less responsive to changes in volatility in terms of the percentage change in the size of their equity portfolio a given volatility movement calls for. However, as a percentage of their total wealth both short and long-term investors respond by identical amounts.

3.3 The (large) costs of ignoring variation in volatility

It is now clear that long-horizon investors should volatility time quite aggressively. Yet one could think that because volatility shocks are not very persistent, it might not be very costly to deviate from the optimal strategy. Here we evaluate the benefits of volatility timing by comparing increases in utility of only using information on conditional volatility, with the fully optimal policy that also uses information on conditional expected returns.

Specifically, we focus on a volatility timing strategies of the form

$$\widetilde{w}(\sigma_t^2) = \omega_0 + \omega_1 w^v(\sigma_t^2), \tag{19}$$

where $w^v(\sigma_t^2) = \frac{\overline{\mu}}{\gamma\sigma_t^2}$ is the weight of a volatility managed portfolio from Moreira and Muir (2016). We refer to $\widetilde{w}(\sigma_t^2)$ as the volatility timing portfolio. We then compute the increase in life-time utility of switching from the myopic buy-and-hold portfolio to a portfolio that holds $\omega_0 = E[w^h]$ in the buy and hold portfolio and $\omega_1 = 1$ in the volatility managed portfolio. We refer to $\widetilde{w}(\sigma_t^2)$ as the optimal volatility timing portfolio[11] We then compare these gains from volatility timing with the utility increase from switching from buy-and-hold to the fully optimal policy.

The first row in Table 3 show results for our point estimates. Following the myopic buy-and-hold strategy is very costly. For the baseline estimates, volatility timing increases utility by 35% relative to buy-and-hold, and switching to the full optimal policy that also uses expected return information increases utility by about 60%. Thus the gains from volatility timing are large when compared to the total benefits of exploiting conditioning information. Specifically, the third column shows that one can capture 60% of the total gains from timing by only timing volatility.

Looking across rows in Table 3 we see that that this pattern holds up across a wide set of parameters. Volatility timing always increases utility relative to buy-and-hold, and typically leads to increases that are more than 50% of the total gains form timing. All comparative statics are very intuitive. Gains make up a larger fraction of total gains when volatility is more variable and more persistent, expected returns less volatile and more persistent, and investment horizons are shorter. Even though as a fraction of the total gains from timing, the gains from volatility timing decreases with the investment horizon, it still accounts for more than 40% of the gains for an investor with a 50 year horizon. Therefore, volatility timing should be an essential component of investors portfolio strategy even when these investors have really long horizons.

Lastly it is worth highlighting the importance of the unconditional expected return.

[11]We do note solve for the optimal weights ω_0 and ω_1, but in unreported results we find that weights $\omega_0 = E[w^h]$ and $\omega_1 = 1$ are very close to the optimal weights.

We see in Table 3 that the gains from volatility timing are strongly increasing in the unconditional expected return. This is consistent with the analysis in Moreira and Muir (2016) who find that when there is no conditional risk-return trade-off, the gains from volatility timing for a myopic investor increase proportionally to the unconditional risk-return trade-off. While we calibrate our the average equity premium to 5%,our sample has an unconditional equity premium average of 7.8%. Thus, the utility gains shown in Table 3 can be thought as conservative estimates.

Overall, these results show that ignoring volatility is likely to be very costly. These costs are large not only for our point estimates but also for a wide range of parameters that are consistent with the data.

3.4 The composition of volatility shocks

We have so far followed the empirical literature and assumed that the composition of volatility shocks is constant (equal to α_μ) (see for example Campbell et al. (2012)). Thus, when volatility changes, discount rate and cash flow volatility change proportionally (this is captured by the conditional volatility of $d\mu_t$ being proportional to σ_t). As a result, the amount of mean-reversion in returns is constant. While this assumption is plausible a priori, it rules out the idea that movements in volatility are mostly due to "short-term volatility". For example, Cochrane (2008a) argued that the huge spike in volatility in the fall of 2008 was fundamentally about an increase in the volatility of transitory shocks:

> "And what about volatility?(...)the standard formula is wrong. (...) Stocks act a lot like long-term bonds – (...)If bond prices go down more, bond yields and long-run returns will rise just enough that you face no long-run risk.(...)the same logic explains why you can ignore "short-run" volatility in stock markets."(Cochrane, 2008a)

Empirically, there is no direct empirical evidence that confirms or refutes Cochrane (2008a) conjecture. In light of the fact that measuring the average share of discount rate shocks is already challenging (Goyal and Welch, 2008), measuring time-variation in the share is even harder. Nevertheless, Golez and Koudijs (2014) provides some indirect evidence that suggests that there might be a positive correlation between volatility and the discount rate share. Using data that goes back to seventeenth century, Golez and Koudijs (2014) show that most of the evidence for return predictability comes from periods identified as economic recessions. Together with the evidence that volatility tends to be high in recessions (Moreira and Muir, 2016; Lustig and Verdelhan, 2012), this evidence suggests that the share of discount-rate volatility might increase in periods of high volatility consistent with the idea that at least some of the movements in volatility are due to short-term volatility.

This correlation matters to an investor because if increases in volatility are entirely due to increases in discount-rate volatility, the increase in return mean-reversion will offset the increase in volatility, making the risky asset just as safe in the long-run. Thus, the intuition is that investors with long investment horizons should not perceive periods of high discount rate volatility as much riskier than low volatility periods.

To capture this idea that volatility variation is driven by discount rate shocks we now explicitly decompose return innovations into discount rate (dB_t^μ) and cash flow shocks (dB_t^c) as

$$dB_t^S = -\sqrt{\frac{\sigma_t^2 - \underline{\sigma}^2}{\sigma_t^2}} dB_t^\mu + \sqrt{\frac{\underline{\sigma}^2}{\sigma_t^2}} dB_t^c, \tag{20}$$

which implies that the volatility of cash flow shocks is constant and only discount rate volatility varies. Consistent with Equation (20) we set the volatility of expected returns in Equation (3) to $\kappa_\mu \sqrt{\sigma_t^2 - \underline{\sigma}^2}$, which implies the discount rate share of return shocks is

$$\frac{\sigma_t^2 - \underline{\sigma}^2}{\sigma_t^2}. \tag{21}$$

This share goes to 1 as volatility spikes to high levels and goes to zero as volatility drifts to the lower bound.[12]

[12] The correlation between cash flow and volatility shocks is simply $< dB_t^c, dB_t^\sigma >= -\rho_{\sigma,S}$

Figure 2 show the IRFs for this extreme case where all volatility variation is due to variation in the volatility of discount rate shocks. In the middle panel we see how the discount rate share spikes up with volatility and then slowly comes down. In the bottom panel we see that this results in an increase in hedging demand, which counteracts the decrease is exposure due to the myopic demand. It is still optimal to reduce the portfolio exposure after a volatility shock, but the response is less aggressive. Intuitively, stocks become relatively safer for a long term investor than for a short-term investor when the share of discount rate shocks goes up. In response to a one standard deviation shock, the investor reduces his position in the risky asset by 25%, substantially less than the 40% in the constant discount rate share case.

The optimal portfolio can still be implemented with a constant position in the buy-and-hold and the myopic portfolio, but now the exposure to the myopic portfolio deviates from 1. In Table 4, we contrast the optimal portfolio and utility gains from timing in a discount rate volatility world with our baseline case where discount rate and cash flow vol go up proportionally with volatility. The general pattern is consistent with Figure 2. The second column shows that a positive co-movement between volatility and the discount rate share implies the optimal portfolio has a lower elasticity ζ. This lower response to volatility also means lower gains from volatility timing. For our point estimates, the utility gain from switching from buy-and-hold to a volatility managed strategy falls from 30% to 20%.

Overall, this section shows that the composition of volatility shocks is a quantitatively important determinant of the optimal response to volatility. The co-movement of volatility with the discount rate share determines whether is optimal to respond more or less aggressively to changes in volatility. Note however that it is always optimal to reduce the position in the risky asset when volatility goes up, and the benefits of such a strategy are always large.[13]

[13] In unreported results we study the case where all volatility is driven by cash flow shocks. We find that in this case the optimal response to volatility variation and utility gains from volatility timing are larger than in our baseline case.

3.5 Why do investors respond to increases in the volatility of purely transitory shocks to returns?

Because in the data shocks are very persistent. Specifically, our SMM estimation interprets the return forecasting R-squares increase from 0.6% at the monthly horizon to 23% at the five year horizon as evidence that the expected return process should be very persistent. This steep slope implies a mean-reversion coefficient of $\kappa_\mu = 0.12$ for movement in expected returns, which translates into an auto-correlation of about 0.88 at the yearly frequency (a half-life of about six years). To put in perspective, an investor with a 20 year horizon (our baseline calibration) consumes about 30% of her wealth during this period. Thus, investors respond to variation in the volatility of discount rate shocks because a substantial fraction of their consumption responds to variation in discount rates.

However, it is well known that return forecasting R-squares based on price-dividend ratios might lead us to over-estimate the persistence of expected returns. Among others, Lettau and Van Nieuwerburgh (2008) and more recently Kelly and Pruitt (2013) find evidence that expected returns are much less persistent than implied the return forecasting R-squares we use in our estimation. For example, Lettau and Van Nieuwerburgh (2008) shows that when they allow for a structural break in the sample, the half-life of expected return shocks decrease to about three years.[14]

Motivated by these findings, here we study the sensitivity for our results to the possibility that that expected returns are less persistent than implied by the moments we match in our estimation. Note however that as the persistence of expected return decrease, the persistence of both processes become more similar (recall that the volatility process is much less persistent than expected returns in our estimation),and the high correlation between expected return and volatility shocks map into a high unconditional correlation, which is inconsistent with the weak risk-return trade-off present in data (see Moreira and Muir (2016)).

In Table 5 we study how our results change as we vary the persistence of the discount rate process κ_μ. We choose values to reflect the wide range of estimates in the literature,

[14] See also evidence presented in Drechsler and Yaron (2011).

which are often found using different methods and predictors. In addition to our baseline result solution with an annualized persistence (auto-correlation) of expected returns equal to 0.89 ($\kappa_\mu = 0.12$), as estimated from price-dividend ratio based return forecasting regressions, we also report persistence to 0.94 ($\kappa_\mu = 0.06$), in line with estimates from the dividend yield coming from Cochrane (2008b), among others, and persistence of 0.78 ($\kappa_\mu = 0.25$). This lower discount rate persistence is more towards the lower end of the persistence of variables that predict returns found in the predictability literature (see e.g., Lettau and Ludvigson (2001), Kelly and Pruitt (2013), or Drechsler and Yaron (2011)).

First looking across the three columns where we vary persistence but keep the standard assumption of constant discount rate share, we see that the elasticity to volatility variation gradually goes down, but it is still above 0.6 even for the very low persistence calibration. We also see that while total gains from timing goes up as the persistence goes down, the gains form volatility timing goes down. Therefore, the fraction of timing gains due to volatility timing goes down sharply from 66% to 39% as as we move from the more to the less persistent calibration. It is important to note that while the R-squares increase more for short horizon forecasts as we reduce the persistence of expected returns, longer horizons also exhibit an increase. This happens here because we setup conditional volatility of expected returns to be $\sigma_\mu = \kappa_\mu \alpha_\mu$ so that the parameter α_μ has the interpretation of a discount rate share. As a result the unconditional volatility of expected returns (approximately $\alpha_\mu \sqrt{\kappa_\mu}$) increases with the parameter κ_μ. Nevertheless this overall increase in volatility of discount rates is convenient to fit the forecasting patterns documented in Kelly and Pruitt (2013) who document much larger R-squares than our estimates for both the short (1 month) and medium range frequencies (1 year).

In the second set of columns we follow the analysis in Section 3.4 and assume that all variation in volatility is driven by variation in the volatility of discount rates. We see now that elasticities go down sharply. Utility gains from volatility timing are as low as 13% with the share of total timing gains now close to 30%. This numbers are still not quite consistent with the "ignoring" volatility advice, but are much closer.

In summary, ignoring volatility variation is less costly to investors if discount rate shocks have a very low persistence and all variation in volatility is about these low persistence transitory shocks. However, even in this case, there are still meaningful benefits to timing volatility.

4. Incorporating Uncertainty

This section does two things. First, we assess the uncertainty surrounding our utility gains. Specifically, rather than only reporting the point estimate for average utility gains, we study the full distribution of utility gains where we use the uncertainty from our estimation procedure about the parameters. We find the gains from volatility timing are extremely likely to be positive. Second, we then incorporate the fact that the parameters are unlikely to be known by the investor ex-ante. We incorporate parameter uncertainty by assuming an investor observes a signal for expected returns and volatility but does not know the true process for each and thus faces estimation risk, along the lines of Barberis (2000).

4.1 What is the probability that ignoring volatility variation is optimal?

The uncertainty surrounding our SMM parameter estimates indicates this probability is zero. We reach this conclusion by leveraging our SMM estimation to recover the uncertainty surrounding our parameter estimates and convert the uncertainty in this estimation to uncertainty about utility gains. We build on our approach to estimate parameter standard errors. Using our point estimates we simulate a sample of identical length as our sample, we then re-estimate the model using the artificial data of the simulation , and use these new estimated parameters to solve for the optimal portfolio choice. We then use the optimal portfolio functions to calculate the elasticity of portfolio with respect to changes in volatility and compute the utility gains from switching from the myopic buy-and-hold strategy to a volatility managed strategy. This approach allow us to recover the full distribution of the optimal portfolio elasticity and the economic gains of volatility timing.

The distribution of these quantities are shown in Figure 3. The portfolio elasticity to volatility ranges from 0.4 to 1. This implies that inaction is never optimal across bootstrapped parameters. Economics gains of volatility timing reflect this large elasticity. We find utility gains of volatility timing range form 20% to 90%. Thus, ignoring volatility variation is extremely likely to be very costly.

It is important to emphasize that here we abstract form uncertainty about the unconditional expected return. As we emphasized in Section 3.3 the level of expected returns is an important determinant of the overall gains of volatility timing. In terms of the average level of utility gains form vol timing, our approach of calibrating the equity risk premium is conservative, but this approach is likely to lead us to underestimate the amount of overall uncertainty in the gains from volatility timing. We plan in the future to incorporate this source of uncertainty in our analysis. We expect the average gains to increase, but to become slightly more uncertain.

4.2 Parameter uncertainty and imperfect information

Our analysis so far endows investors' with perfect information with respect to variation in the investment opportunity set, i.e. we take our in-sample point estimates for each parameter as the true generating process and assume this is known to the investor. In practice investors have to form portfolios and trading strategies while facing uncertainty about the true return generating process.

We now consider the case where the investor is not sure of the process for returns and thinks about parameter uncertainty. We investigate how sensitive the utility gains of volatility timing and expected return timing are to the parameter uncertainty present in the data. Our goal is to assess how this separately affects the utility gains from volatility timing and expected return timing described before. This is important because, as many papers have shown, parameter uncertainty surrounding return predictability can have very large effects (Barberis (2000), Goyal and Welch (2008), Cochrane (2008b), Pástor and Stambaugh (2012)). We confirm these results but show that parameter uncertainty is not a big issue for volatility timing. Our approach to parameter uncertainty follows closely the method described in Barberis (2000).

Specifically, we take the approach of an investor who is given a 90 year sample for returns and who must estimate a rule to forecast returns and to forecast volatility using this 90 year sample. The investor then adopts permanently a portfolio timing rule based on this forecasting relationship. We then compute the expected utility for the investor going forward and compare this expected utility to the case where the investor knows the true process for returns (as computed earlier).

We begin by assessing the benefits of expected return timing given this parameter uncertainty, which we implement as follows. In a given 90 year sample, the investor observes the variable x_t which is the true conditional expected return for the risky-asset, and uses this variable to forecast returns in a given sample.[15] One can think of this as running predictive regressions with a candidate predictor such as the price dividend ratio. The regression the investor runs in each sample is

$$r_{t+1} = a + bx_t + \varepsilon_{t+1}, \quad t = 1, ..., T$$

The true value for this regression is $a = 0$, $b = 1$, but the investor does not know this, he only sees x_t as a candidate predictor of returns. In estimating this regression using a given sample, s, the investor estimates $\hat{a}_s, \hat{b}_s$ where these are the coefficients recovered in a given sample. He then devises a trading strategy for what he believe is the expected return process going forward using these coefficients as the fitted value from this regression $\hat{\mu}_{T+t,s} = \hat{a}_s + \hat{b}_s x_{T+t}$ applied to the optimal portfolio rule $\widetilde{w}(\mu_t, \sigma_t^2)$ described earlier in the paper. Notice that if, given 90 years of data, the investor always recovered the true coefficients a and b, then this would be equivalent to the utility benefits of the full timing case studied earlier. We assume the investor knows all other parameters of the return process so as to isolate only the effects coming from not knowing expected returns.

We then ask what is the expected utility associated with this rule given that these coefficients may vary from sample to sample, i.e., given that, even with 90 years of data, the investor may not know the true relationship between the predictor variable and future returns? These results are given in the bottom of Table 6. It turns out that the expected utility of the investor is much lower than the case when he knows the true expected re-

[15] Note that the estimation uncertainty we document here would be even larger if we were to consider imperfect predictors, that is, the investor only observes a noisy signal of x_t, see Pástor and Stambaugh (2012).

turn. The reason is that these estimated coefficients $\hat{a}_s, \hat{b}_s$ vary dramatically from sample to sample, even with 90 years of data. Largely this has to do with expected returns being very persistent making the relationship in the predictability regression difficult to estimate in a given sample. This point is well recognized by Goyal and Welch (2008), among others.

Next, we consider that the investor needs to make a forecast for volatility. Here, we assume the investor observes realized volatility in a given sample, where realized volatility is the volatility of daily returns in a given month. He uses this realized volatility to try to forecast the process for volatility in the next month – that is the investor maps realized volatility to expected volatility in a 90 year sample, and then uses this going forward in the following periods to forecast volatility. This is captured by $(r_{t+1} - \mu_t)^2 = a + b\sigma_t^2 + \varepsilon_{t+1}$ where we study the variation in these coefficients as before. As was the case before with the return forecasting regression, the investor is given a perfect signal, σ_t, about volatility, but must use this signal in the given sample to make a forecast about future volatility. Thus, the investor may poorly estimate the relationship between realized volatility and true volatility – analogous to the difficultly in predicting expected returns. This turns out to be inconsequential – given 90 years of data, and the much lower persistence of volatility, the investor faces very low estimation risk. Table 6 contains these results and shows that the utility gains for volatility timing are essentially preserved when we take into account estimation risk.

In both the return and volatility forecasts we assumed the investor had perfect signals of the true process – in practice this is much more likely to be true of volatility as investors observe realized volatility and have signals like the VIX which give near perfect signals of volatility in real time. In contrast, it is less likely that the investor would have a perfect signal for expected returns. Thus, our analysis here if anything understates the affects of parameter uncertainty on expected returns if one incorporates imperfect predictors.

In summary, because expected returns appear very persistent, predictive variables in a given sample can work poorly as forecasts for returns out of sample. This means that the benefits of timing expected returns are very sensitive to parameter uncertainty. We confirm this fact here, but this fact is well documented. However, this result is not true with volatility – this is essentially because volatility is easy to forecast both in and out of sample. Hence, the utility gains from volatility timing are far more robust to parameter uncertainty.

5. Extensions

We consider a number of extensions to our model including alternative preferences, and outside income risk and we briefly discuss how these extensions might interact with volatility timing.

5.1 Alternative preferences

Thus far we have studied EZ preferences with unit elasticity of substitution. These preference are not only standard in the portfolio choice literature, but also very convenient as we can directly control the investor horizon by varying the impatience parameter ρ. Alternative preferences studied in the literature include: (1) EZ preferences with non-unit IES, (2) Constant Relative Risk Aversion preferences (i.e., IES=1/RRA), and (3) preferences with habit formation.

Here we extend our analysis to (1) and (2). Figure 5 show these results by comparing impulse responses across preferences. We normalize each portfolio weight by it's steady state value so we can focus exclusively of the weight elasticity to a volatility shock. Panel A shows the volatility IRF and Panels B to D show the portfolio response in three different cases. Starting with Panel B, which shows results for the baseline assumption that the composition of volatility shocks is constant we see that all investors respond identically. The same is true for Panel D, the case of negative co-movement between the discount rate share and volatility. Only in the case of positive co-movement we see some differences across investors responses. Most interesting we see that high IES investors tend to respond less to a volatility shock in this case. The reason behind this result is intuitive. Investors with higher IES tend to be less responsive to discount rate volatility because they optimally choose to save more when the investment opportunity set is very attractive, i.e. because their are more willing to postpone consumption their horizon is endogenously longer when the opportunity set is more attractive.

Case (3), habit, is substantially more complicated, as it requires adding a habit state variable. We haven't analyzed this case explicitly, but the analysis in Detemple and Zapatero (1992) and Gomes and Michaelides (2003) suggests that such preferences will lead us to similar results. For example, Detemple and Zapatero (1992) show that habit formation leads investors to first invest in a perfectly safe portfolio that finances habit consumption, and then invest as a standard CRRA agent that had only the residual wealth (the wealth minus the safe portfolio) would. Thus, they respond to a volatility shock as a CRRA agent with a similar allocation to the risky asset would. Their result suggest that while agents with habit forming preferences will invest much less in the market, their elasticity to a volatility shock is equal to a standard CRRA investor.

5.2 Non-financial income

Our baseline analysis is purposefully stark as it relies on the assumption that the investor only source of income is her financial wealth. A more realistic assumption is that the investor also earns wage or other sources of income. For example, Merton (1971), Viceira (2001), Cocco et al. (2005) , and Polkovnichenko (2007) are examples of recent work that study how non-financial income shape portfolio decisions. For the baseline case where outside income is risk-less, these papers show that optimal portfolio is simply

$$w_t \frac{W_t}{W_t + PDV_t(E)} = \frac{\mu_t}{\gamma\sigma_t^2} + \text{hedging demand}, \tag{22}$$

where $PDV_t(E)$ is the present discounted value of the investor non-financial income, W_t is the investor financial wealth, and w_t is the share of financial invested in the risky asset. The solution implies that the investor targets the same share of total wealth allocated to the risky asset, what implies a much higher share of financial wealth, as $\frac{W_t}{W_t+PDV_t(E)} < 1$. In this simple risk-less case, the solution is analogous to the investor having a lower risk-aversion, $\tilde{\gamma} = \frac{W_t}{W_t+PDV_t(E)}\gamma$. Thus, all our our result will carry through to this case. We simply need to use $\tilde{\gamma}$ as the investor coefficient of relative risk-aversion.

The impact of risk in the non-financial income stream can be understood by decomposing it in an idiosyncratic component that cannot be hedged or diversified, and a component that co-varies with the risky asset. Both idiosyncratic has the effect of reducing the present discounted value of the outside income. Intuitively, the higher the volatility, the higher the co-variance between the income and the investor marginal utility. The end result is a higher discount-rate. Again, our results will apply as in the risk-less case after adjusting the outside income present discounted value. The aggregate component has a second effect because it not only impacts the value of the income stream, but because it can be hedged. The effect on the present value of the income stream is straightforward: a positive exposure increases the discount rate according to the risk-premium earned in the risky asset. The co-variance with the risky asset induces a new kind of hedging demand to emerge. Intuitively, the optimal portfolio choice adjusts for any exposure the investors income already has to the risky asset. A positive co-variance thus induces a negative hedging demand, reducing the share of the investor financial wealth allocated to stocks.

While in practice, it is hard to find industries with wage income that is sufficiently strongly correlated with the stock market for these hedging demands to matter, more sophisticated modeling of labor income risk emphasizes a long run relation between the stock market and wages. For example, Benzoni et al. (2007) show that if labor income is co-integrated with dividends, the hedging demand can be large for empirically plausible parameters. Could this type of hedging demand overturn our results? As we have seen in Section 3.2, a constant negative hedging demand has the effect of increasing the elasticity of the portfolio weight to volatility. Thus, the "level" of the hedging demand will tend to amplify the optimal response to volatility. Our results can be overturn only if outside income hedging demand increases with volatility (push the portfolio towards stocks). The logic of co-integration is that all permanent shocks to stock prices, i.e. cash-flow shocks, end up eventually impacting the labor income. Thus variation in cash-flow volatility should translate one-to-one to variation in the hedging demands, i.e. the hedging demand should become more negative in response to an increase in volatility. Variation in discount rate volatility on the other hand would not impact the hedging demand in this case. Thus, this co-integration channel would either increase or not impact the portfolio elasticity to volatility.

In order for the hedging demand to actually go up as volatility increases, the correlation between stock returns and wage income would have to go down enough to more than off-set the increase in volatility. That is, a constant co-variance between wage income and stock returns is not sufficient to overturn our results. In fact, this co-variance would have to be strongly negatively related to volatility in order for hedging demand to increase with volatility. We don't know of any empirical evidence pointing in this direction.

6. Conclusion

We study the portfolio problem of a long-lived investor that allocates her wealth between a risk-less and a risky asset in an environment where both volatility and expected returns are time-varying. We then comprehensively and quantitatively study how investors should respond to changes in volatility and what the utility costs to ignoring volatility variation are, and we study how these results change with the investor's horizon. Importantly, our analysis also takes into account that investors' face parameter uncertainty regarding the dynamics of volatility and expected returns.

The main finding in this paper is that investors should substantially *decrease* risk exposure after an increase in volatility and that ignoring variation in volatility leads to large utility losses. The benefits of volatility timing are on the order of 50% of lifetime utility for our preferred parameterization of an investor with risk aversion of 5 and a 50 year horizon. These benefits are significantly larger than those coming from expected return timing (i.e., from return predictability), particularly when parameter uncertainty is taken into account. We approximate the optimal volatility timing portfolio and find that its dependence on volatility is very simple: all investors, regardless of horizon, will choose fixed weights on a buy-and-hold portfolio that invests a constant amount in the risky-asset, and a volatility managed portfolio that scales the risky-asset exposure by the inverse of expected variance $1/\sigma_t^2$. Further, we show that the weight on the volatility timing portfolio is independent of the investors' horizon in our baseline results.

We then show a novel channel through which long-horizon investors may differ in their response to volatility: they respond less aggressively to increases in volatility when only the volatility of mean-reverting shocks increases. Intuitively, this effect makes stock prices more volatile in the short run but doesn't change the distribution of long run stock prices. This effect can dampen, but does not eliminate, long horizon investors' response to changes in volatility.

References

Barberis, N. (2000). Investing for the long run when returns are predictable. *The Journal of Finance*, 55(1):pp. 225–264.

Benzoni, L., Collin-Dufresne, P., and Goldstein, R. S. (2007). Portfolio choice over the life-cycle when the stock and labor markets are cointegrated. *The Journal of Finance*, 62(5):2123–2167.

Brandt, M. W. (1999). Estimating portfolio and consumption choice: A conditional euler equations approach. *The Journal of Finance*, 54(5):1609–1645.

Buffett, W. E. (2008). Buy American. I am. *The New York Times*, October 16th, 2008.

Buraschi, A., Porchia, P., and Trojani, F. (2010). Correlation risk and optimal portfolio choice. *The Journal of Finance*, 65(1):393–420.

Campbell, J. Y. (1996). Understanding risk and return. *Journal of Political Economy*, 104(2):572–621.

Campbell, J. Y., Giglio, S., Polk, C., and Turley, R. (2012). An intertemporal capm with stochastic volatility. Technical report, National Bureau of Economic Research.

Campbell, J. Y. and Shiller, R. J. (1988). The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies*, 1(3):195–228.

Campbell, J. Y. and Viceira, L. M. (1999). Consumption and portfolio decisions when expected returns are time varying. *Quarterly Journal of Economics*, 114(2).

Campbell, J. Y. and Vuolteenaho, T. (2004). Bad beta, good beta. *American Economic Review*, pages 1249–1275.

Chacko, G. and Viceira, L. M. (2005). Dynamic consumption and portfolio choice with stochastic volatility in incomplete markets. *Review of Financial Studies*, 18(4):1369–1402.

Cocco, J. F., Gomes, F. J., and Maenhout, P. J. (2005). Consumption and portfolio choice over the life cycle. *Review of Financial Studies*, 18(2):491–533.

Cochrane, J. H. (2008a). Is now the time to buy stocks? *The Wall Street Journal*, November 12th, 2008.

Cochrane, J. H. (2008b). The Dog That Did Not Bark: A Defense of Return Predictability. *Review of Financial Studies*, 21(4):1533–1575.

Collin-Dufresne, P. and Lochstoer, L. (2016). Long-horizon investment, mean reversion, and structural breaks: A general equilibrium perspective. *working paper*.

Cox, J. C., Ingersoll, J. E., and Ross, S. A. (1985). An intertemporal general equilibrium model of asset prices. *Econometrica*, 53(2):363–384.

Detemple, J. B. and Zapatero, F. (1992). Optimal consumption-portfolio policies with habit formation1. *Mathematical Finance*, 2(4):251–274.

Drechsler, I. and Yaron, A. (2011). What's vol got to do with it. *Review of Financial Studies*, 24(1):1–45.

Duffie, D. and Epstein, L. G. (1992). Stochastic differential utility. *Econometrica: Journal of the Econometric Society*, pages 353–394.

Duffie, D. and Singleton, K. J. (1993). Simulated moments estimation of markov models of asset prices. *Econometrica*, 61(4):929–952.

Epstein, L. G. and Zin, S. E. (1989). Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica: Journal of the Econometric Society*, pages 937–969.

French, K. R., Schwert, G. W., and Stambaugh, R. F. (1987). Expected stock returns and volatility. *Journal of financial Economics*, 19(1):3–29.

Glosten, L. R., Jagannathan, R., and Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The journal of finance*, 48(5):1779–1801.

Golez, B. and Koudijs, P. (2014). Four centuries of return predictability. Technical report, National Bureau of Economic Research.

Gomes, F. and Michaelides, A. (2003). Portfolio choice with internal habit formation: A life-cycle model with uninsurable labor income risk. *Review of Economic Dynamics*, 6(4):729–766.

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

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of financial studies*, 6(2):327–343.

Johannes, M., Korteweg, A., and Polson, N. (2014). Sequential learning, predictability, and optimal portfolio returns. *The Journal of Finance*, 69(2):611–644.

Kelly, B. and Pruitt, S. (2013). Market expectations in the cross-section of present values. *The Journal of Finance*, 68(5):1721–1756.

Lettau, M. and Ludvigson, S. (2001). Consumption, aggregate wealth, and expected stock returns. *the Journal of Finance*, 56(3):815–849.

Lettau, M. and Ludvigson, S. (2003). Measuring and modelling variation in the risk–return tradeoff, prepared for the handbook of financial econometrics, edited by y. *Ait–Sahalia and Lars–Peter Hansen*.

Lettau, M. and Van Nieuwerburgh, S. (2008). Reconciling the return predictability evidence. *Review of Financial Studies*, 21(4):1607–1652.

Liu, J. (2007). Portfolio selection in stochastic environments. *Review of Financial Studies*, 20(1):1–39.

Lustig, H. and Verdelhan, A. (2012). Business cycle variation in the risk-return trade-off. *Journal of Monetary Economics*, 59, Supplement(0):S35 – S49.

Merton, R. C. (1971). Optimum consumption and portfolio rules in a continuous-time model. *Journal of economic theory*, 3(4):373–413.

Moreira, A. and Muir, T. (2016). Volatility managed portfolios. *Journal of Finance*, forthcoming.

Pástor, L. and Stambaugh, R. F. (2012). Are stocks really less volatile in the long run? *The Journal of Finance*, 67(2):431–478.

Polkovnichenko, V. (2007). Life-cycle portfolio choice with additive habit formation preferences and uninsurable labor income risk. *Review of Financial Studies*, 20(1):83–124.

Poterba, J. M. and Summers, L. H. (1988). Mean reversion in stock prices: Evidence and implications. *Journal of financial economics*, 22(1):27–59.

Samuelson, P. A. (1969). Lifetime portfolio selection by dynamic stochastic programming. *The review of economics and statistics*, pages 239–246.

Schwert, G. W. (1989). Why does stock market volatility change over time? *The journal of finance*, 44(5):1115–1153.

Viceira, L. M. (2001). Optimal portfolio choice for long-horizon investors with nontradable labor income. *The Journal of Finance*, 56(2):433–470.

Wachter, J. A. (2002). Portfolio and consumption decisions under mean-reverting returns: An exact solution for complete markets. *Journal of financial and quantitative analysis*, 37(01):63–91.

Xia, Y. (2001). Learning about predictability: The effects of parameter uncertainty on dynamic asset allocation. *The Journal of Finance*, 56(1):205–246.

7. Tables and Figures

Table 1: Estimates for the stochastic processes.Panel A provides the parameters we calibrate and their calibrated values. Panel B the parameters we estimate together with their standard errors. Panel C provides the matched moments in the data and in the model, together with their standard errors. Parameters are estimated using simulated method of moments (SMM). We report bootstrapped standard errors. For the moments this consists of using the model to simulate 1000 samples of equal size as ours and then computing the standard error of each moments across samples. For the parameters this consists of re-estimating the model for each of these sample realizations and then computing the standard errors across point the point estimates of each simulated sample. Note that $RV_t = \sum_{d=1/22}^{1}(R^e_{t+d} - (\sum_{d=1/22}^{1} R^e_{t+d}/22))$ is the daily realized variance in month t, where R^e_{t+d} is the excess return on the market on date $t+d$. Small cap $rv_t = \ln(RV_t)$ is short for log realized variance. $R^e_{t\to T}$ is the excess return on the market between dates t and T. The alpha of the volatility managed portfolio $\alpha(SMKT \to MKT)$ is the intercept of a regression of the volatility managed (excess) market portfolio on the (excess) market portfolio itself (See Moreira and Muir (2016) for details).

Panel A: Calibrated parameters

Parameter	Description	Value
r	Risk-free rate	0.01
$\overline{\mu}$	Equity premium	0.05
$\underline{\sigma}$	Volatility lower bound	0.07

Panel B: Estimated parameters

Parameter	Description	Point estimate	s.e.	10%	90%
$\overline{\sigma}$	Avg vol	0.17	0.02	0.14	0.20
κ_σ	Vol persistence	3.34	0.45	2.70	3.83
ν_σ	Vol volatility	5.25	0.40	4.89	5.82
α_μ	Share of discount rate shocks	0.59	0.34	0.02	1.00
κ_μ	Discount rate shocks persistence	0.12	0.01	0.12	0.14
$\rho_{\sigma,S}$	Corr vol and CF shocks	0.44	0.31	0.04	0.95
$\rho_{\sigma,\mu}$	Corr vol and DR shocks	0.57	0.19	0.26	0.82

Panel C: Estimated Moments

Moment	Description	Data	Model	std. er.	10%	90%
$\sqrt{E(RV_t)}$	Avg. realized vol. (annual units)	16.787%	16.756%	2.090	14.296	19.688
$\rho(rv_t, rv_{t-1})$	Auto-corr. of log vol.	0.722	0.704	0.029	0.672	0.743
$stdev(rv_t - \widehat{rv}_t)$	Std dev of log vol. shocks	0.729	0.715	0.033	0.671	0.757
$corr(rv_t - \widehat{rv}_t, R_t)$	Corr(vol. shocks, returns)	-0.360	-0.344	0.046	-0.410	-0.289
$R^2(R^e_{t\to t+1} \to pd_t)$	Predict. regression R-sq (1 month)	0.591%	0.585%	0.418	0.096	1.251
$R^2(R^e_{t\to t+60} \to pd_t)$	Predict. regression R-sq (5 years)	23.427%	21.740%	14.224	3.524	40.685
$\alpha(SMKT \to MKT)$	Alpha of vol. managed portfolio	4.830 %	4.446%	2.435	1.068	7.364

Table 2: Optimal portfolio. We first show ζ, the local elasticity of the optimal portfolio to changes in variance $\zeta = -dlog(w_t^*)/dlog(\sigma_t^2)$. We then show the approximation of the optimal portfolio that is affine in the myopic portfolio $\widetilde{w}^*(\mu_t, \sigma_t^2) = \omega_0 + \omega_1 w^m(\mu_t, \sigma_t^2)$ where $w^m(\mu_t, \sigma_t^2) = \frac{\mu_t}{\gamma\sigma_t^2}$ and γ is investor risk aversion. We report weights in the static buy-and-hold portfolio (ω_0) and the myopic portfolio (ω_1). The last column computes the utility losses from following the affine portfolio $\widetilde{w}$ compared to the optimal portfolio, i.e. $\Delta\widetilde{U}^* = U[\widetilde{w}^*]/U[w^*] - 1$. It shows that our linear approximation captures the true optimal portfolio well in terms of resulting in small utility losses. Utility losses are in wealth units (e.g. a 1% loss is equivalent to a 1% state-by-state loss in the investor life-time consumption)

Parameter	value	ζ	ω_0	ω_1	R^2	$\Delta\widetilde{U}^*$
baseline		0.68	0.27	1.00	0.99	-0.00
$1/h$	10	0.72	0.20	1.00	0.99	-0.00
	50	0.65	0.33	1.00	0.99	-0.00
γ	3	0.70	0.38	1.00	1.00	-0.00
	10	0.66	0.15	1.00	0.99	-0.00
κ_σ	2.67	0.71	0.26	1.00	0.99	-0.01
	3.82	0.67	0.27	1.00	1.00	-0.00
ν_σ	4.94	0.68	0.26	1.00	0.99	-0.00
	5.83	0.68	0.28	1.00	0.99	-0.00
α_μ	0.02	1.03	-0.04	1.00	1.00	-0.00
	1.00	0.37	1.01	1.00	0.99	-0.03
κ_μ	0.12	0.68	0.27	1.00	0.99	-0.00
	0.14	0.67	0.28	1.00	0.99	-0.00
$\rho_{S,\sigma}$	0.04	0.63	0.35	1.00	1.00	-0.00
	0.62	0.67	0.26	1.00	0.99	-0.00
$\rho_{\mu,\sigma}$	0.25	0.64	0.29	1.00	1.00	-0.03
	0.65	0.68	0.28	1.00	0.99	-0.00
$\overline{\mu}$	0.03	0.59	0.22	1.00	1.00	-0.00
	0.07	0.74	0.31	1.00	0.99	-0.00

Table 3: Utility gains from timing. We compare utility gains from several portfolio strategies. The first strategy is the optimal linear portfolio $\widetilde{w}^*(\mu_t, \sigma_t^2) = \omega_0 + \omega_1 w^m(\mu_t, \sigma_t^2)$ (with associated utility $U[\widetilde{w}^*]$) where $w^m(\mu_t, \sigma_t^2) = \frac{\mu_t}{\gamma\sigma_t^2}$ is the myopic portfolio weight. The second, $\widetilde{w}(\sigma_t^2) = \omega_0 + \omega_1 w^v(\sigma_t^2)$ is the volatility timing portfolio, which is the approximation of the optimal portfolio that is an affine function of the volatility managed portfolio $w_t^v = \frac{\overline{\mu}}{\gamma\sigma_t^2}$. The third is the myopic buy-and-hold portfolio $\bar{w} = \overline{\mu}/(\gamma\overline{\sigma}^2)$. The first column shows the the utility gain from switching from buy-and-hold to the optimal linear portfolio ($\Delta U^* = U(\widetilde{w}^*)/U(\bar{w}) - 1$), the second column shows the utility gains from switching from buy-and-hold to the volatility timing portfolio ($\Delta U^\sigma = U(\widetilde{w}[\sigma_t^2])/U(\bar{w}) - 1$). The third columns shows the faction of the total utility gain from switching to the optimal portfolio can be achieved with the volatility timing portfolio ($\Delta U^\sigma/\Delta U^*$) Utility gains are in wealth units (e.g. a 1% gain is equivalent to a 1% state-by-state increase in the investor life-time consumption) .

Parameter	value	ΔU^*	ΔU^σ	$\Delta U^\sigma/\Delta U^*$
baseline		57	34	60
$1/h$	10	31	22	70
	50	184	77	42
γ	3	112	63	56
	10	33	23	70
κ_σ	2.67	68	49	72
	3.82	54	31	58
ν_σ	4.94	54	32	59
	5.83	65	43	65
α_μ	0.02	80	75	93
	1.00	154	95	62
κ_μ	0.12	57	34	59
	0.14	58	32	56
$\rho_{S,\sigma}$	0.04	61	37	60
	0.62	58	32	54
$\rho_{\mu,\sigma}$	0.25	56	30	53
	0.65	57	35	61
$\overline{\mu}$	0.03	40	21	52
	0.07	117	75	63

Table 4: Time-varying composition of volatility shocks. Here we show how time-variation in the discount rate share impacts the optimal investment strategy. In the first column this share is constant (our baseline case), and in the next column the discount rate share is increasing with volatility, which we label "Increasing share." In short, while the degree of mean-reversion in returns is constant in our baseline case, in the second column there is relatively more mean-reversion in returns when volatility is high. Specifically, the second column sets the volatility of "permanent" shocks to returns, which we label "cash flow shocks", to a constant equal to 7%. Thus, in this case, all variation in return volatility is about the volatility of discount rate shocks, or transitory shocks, to returns. When volatility is low, returns are mainly driven by permanent (cash flow) shocks, making stock returns risky for long run investors. An increase in volatility in this case is associated with a large increase in short run risk in returns but less of an increase in long run risk in returns because only the volatility of transitory shocks increases. We show the weights ω_0 and ω_1 that implement the optimal linear portfolio $\widetilde{w}^*(\mu_t, \sigma_t^2) = \omega_0 + \omega_1 w^m(\mu_t, \sigma_t^2)$ where $w^m(\mu_t, \sigma_t^2) = \frac{\mu_t}{\gamma\sigma_t^2}$, and we compare the utility from alternative portfolio strategies. The first, $\widetilde{w}^*(\mu_t, \sigma_t^2)$ is the optimal linear portfolio (with associated utility $U[\widetilde{w}^*]$). The second, $\widetilde{w}(\sigma_t^2)$ is the volatility timing portfolio. It is an approximation of the optimal portfolio that is affine in the volatility managed portfolio $w^\sigma(\sigma_t^2)$. The third is the myopic buy-and-hold portfolio $\bar{w} = \overline{\mu}/(\gamma\overline{\sigma}^2)$. The row denoted ΔU^* shows the utility gain for an investor going from the myopic buy-and-hold portfolio to the optimal linear portfolio ($\Delta U^* = U[\widetilde{w}^*]/U[\bar{w}] - 1$). The row denoted ΔU^σ shows the utility gain for an investor going from the buy-and-hold portfolio to the volatility timing portfolio ($\Delta U^\sigma = U[\widetilde{w}(\sigma_t^2)]/U[\bar{w}] - 1$). The last row shows the fraction of the total utility gain from the optimal portfolio can be achieved with the volatility timing portfolio ($\Delta U^\sigma/\Delta U^*$). Utility gains are in wealth units (e.g. a 1% gain is equivalent to a 1% state-by-state increase in the investor life-time consumption).

	Constant share	Increasing share
ζ	0.68	0.33
ω_0	0.24	0.43
ω_1	1	0.87
ΔU^*	57	50
ΔU^σ	34	21
$\Delta U^\sigma/\Delta U^*$	60	42

Table 5: The importance of discount rate persistence. We compare the optimal response and utility gains of volatility timing when we change the persistence of discount rates (expected returns). Each column refers to a different set of parameters: the second column shows our baseline solution with an annualized persistence (auto-correlation) of expected returns equal to 0.89 ($\kappa_\mu = 0.12$), as estimated from the data. The first column increases this persistence to 0.94, in line with estimates from the dividend yield coming from Cochrane (2008b), among others. The third column lowers this persistence down to 0.78 ($\kappa_\mu = 0.25$). This lower discount rate persistence is more towards the lower end of the persistence of variables that predict returns found in the predictability literature (see e.g., Lettau and Ludvigson (2001), Kelly and Pruitt (2013), or Drechsler and Yaron (2011)). In the first three columns the discount rate share is constant, and in the last three the discount rate share increases with volatility ("Increasing share"). The first three rows describe the optimal policy in terms of the local elasticity ζ and the weights ω_0 and ω_1 that implement the optimal linear portfolio $\widetilde{w}^*$. The row denoted ΔU^* shows the utility gains from switching from a naive buy-and-hold portfolio ($\overline{w} = \overline{\mu}/\gamma\overline{\sigma}^2$) to the optimal linear portfolio ($\Delta U^* = U[\widetilde{w}^*]/U[\overline{w}] - 1$). The row denoted ΔU^σ shows the utility gain for an investor going from the buy-and-hold portfolio to the volatility timing portfolio ($\Delta U^\sigma = U[\widetilde{w}(\sigma_t^2)]/U[\overline{w}] - 1$). The last row shows the faction of the total utility gain from switching to the optimal linear portfolio can be achieved with the volatility timing portfolio ($\Delta U^\sigma/\Delta U^*$). Utility gains are in wealth units (e.g. a 1% gain is equivalent to a 1% state-by-state increase in the investor life-time consumption).

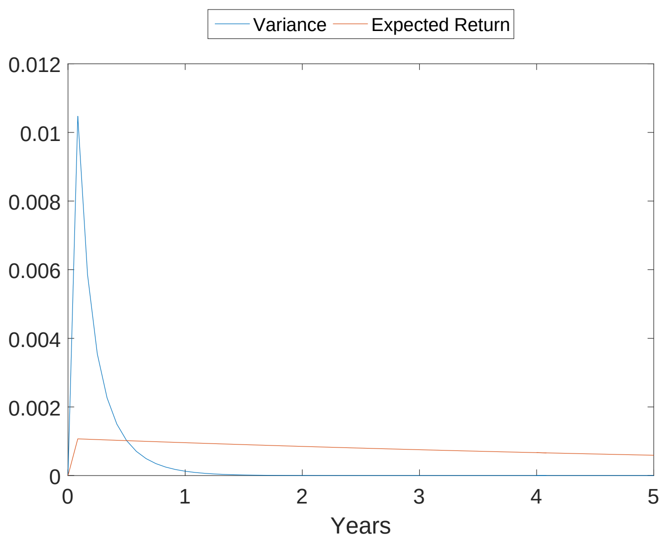
	Constant share			Increasing share		
$\exp(-\kappa_\mu)$	0.94	0.89	0.78	0.94	0.89	0.78
ζ	0.70	0.68	0.64	0.43	0.33	0.29
ω_0	0.21	0.24	0.27	0.32	0.43	0.45
ω_1	1.02	1.03	1.05	0.89	0.87	0.86
ΔU	56	57	61	42	50	39
ΔU^σ	37	34	24	21	21	13
$\Delta U^\sigma/\Delta U$	66	60	39	49	42	33
R-sq(1 month)	0.38	0.56	0.97	0.36	0.73	1.28
R-sq(1 year)	4.25	6.19	9.84	4.40	7.42	12.24
R-sq(5 years)	17.05	21.23	24.32	11.58	25.63	32.26

Table 6: Imperfect information and the costs of estimation uncertainty. This Table evaluates the robustness of adopting a volatility timing with respect to parameter uncertainty. Specifically, we evaluate the utility costs when the investors must use a 90 year sample to estimate a forecasting model for expected returns and volatility. Section 4.2 describes the calculation in detail. The first row shows utility gains when the investor faces no estimation uncertainty and now the expected return and volatility signals. The second row shows utility gains once estimation uncertainty is factored in. Specifically, we report the average utility gain across Bootstrapped estimation samples. The last rows show the distribution of expected utility gains across estimation samples.

	Optimal timing	Volatility Timing
Perfect information	57	34
Imperfect information	5	24
Distribution of expected utility gains		
5%	-90	-54
25%	-16	28
50%	26	33
75%	42	38
95%	50	41

Figure 1: Optimal portfolio response to a volatility shock. The left panel shows the behavior of the conditional expected return and conditional variance after a volatility shock. The right panel shows the optimal portfolio response to a volatility shock for both a mean-variance (short horizon) and a long horizon investor (labeled optimal). It also plots the conditional hedging demand term of the long horizon investor. The top panel does this for the case of a pure volatility shock (so that expected returns do not move). The bottom panel allows a correlation between volatility and expected returns estimated from the data. X-axis is in years.

Panel A: Volatility shock



Panel B: Portfolio response

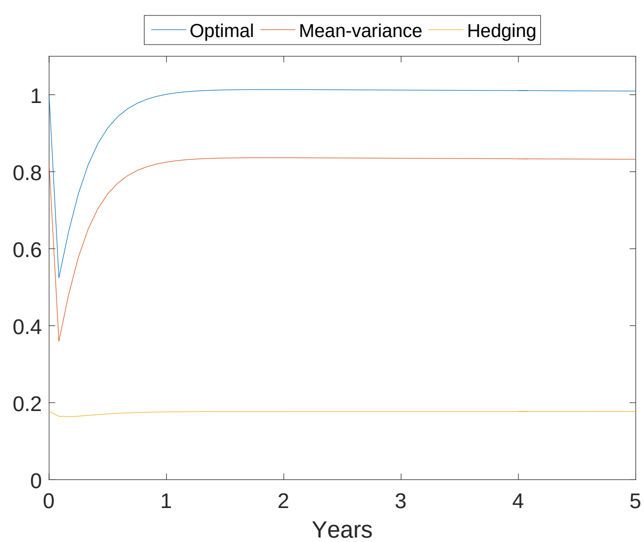


Figure 2: The composition of volatility shocks and the optimal portfolio response to a volatility shock. We repeat the exercise in the previous figure but now allow the composition of volatility shocks to change. Specifically, we now assume that the entire variation in volatility is due variation in the volatility of discount rate shocks. Thus, the share of discount rate shocks in this case is increasing with volatility. To implement this, we assume the volatility of permanent shocks to returns (which we label "cash flow shocks," consistent with the literature) is constant equal to 7% and all time variation in volatility corresponds to discount rate volatility. This generates more mean-reversion in returns when volatility spikes, and hence increases the hedging demand term in high volatility periods as shown in the lower right figure. See section 3.4 for a detailed description. The top plot show the response of volatility and expected return to a one standard deviation volatility shock, the bottom left panel shows how the share of discount rate shocks increases, and the bottom right panel shows the optimal portfolio response.

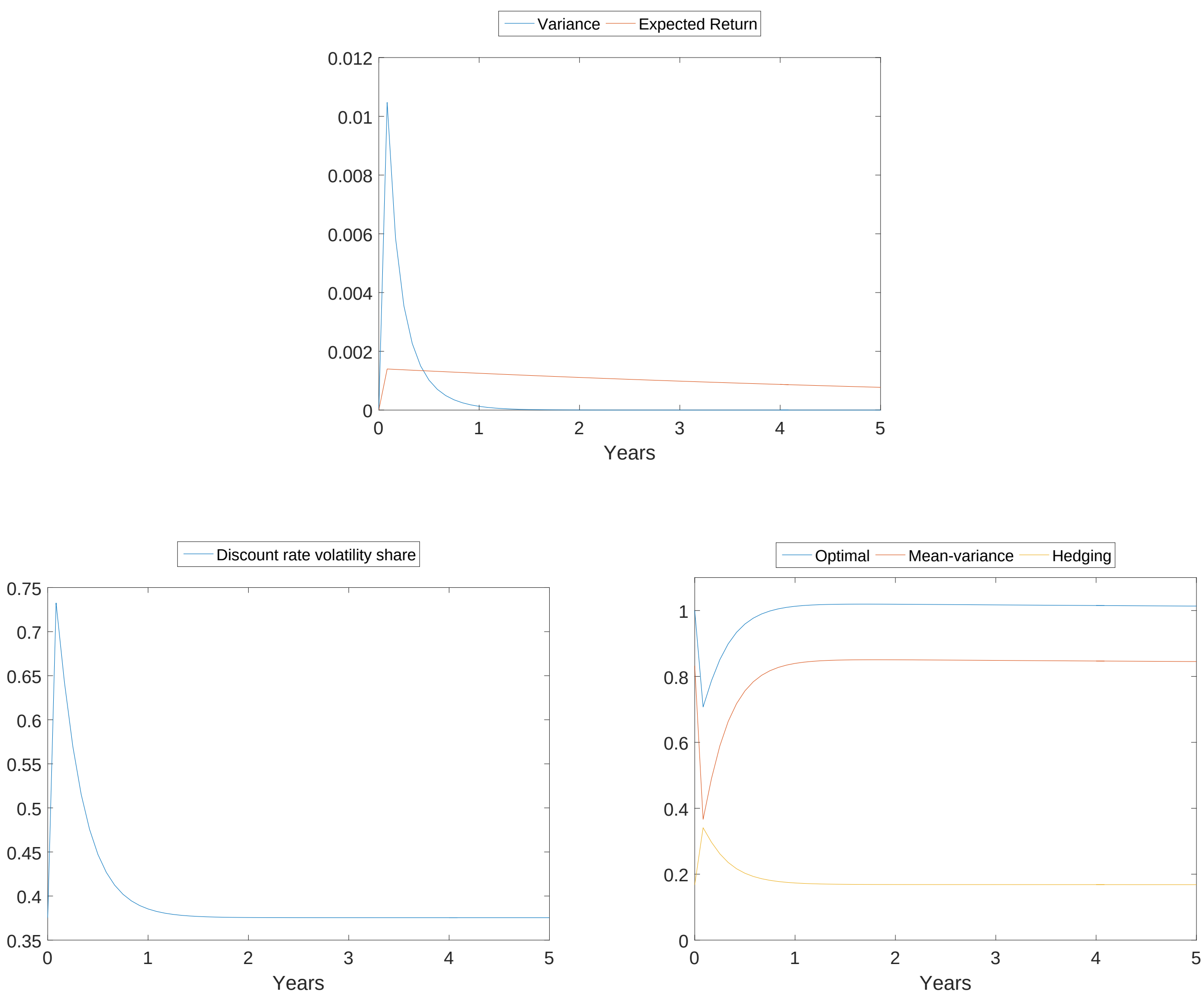


Figure 3: Estimation uncertainty. This Figure shows how estimation uncertainty impacts our results. We use our SMM estimation and optimal portfolio solution to recover the cumulative distribution of several feature of the optimal portfolio policy. The top left Panel shows ζ, the local elasticity of the optimal portfolio to changes in variance $\zeta = -\partial log(w^*)/\partial log(\sigma^2)$. The top right Panel shows the optimal linear portfolio weight in the static buy and hold portfolio (ω_0). The bottom left Panel shows the utility gain for an investor going from the buy-and-hold portfolio to the optimal volatility timing portfolio ($\Delta U^\sigma = U[\widetilde{w}(\sigma_t^2)]/U[\bar{w}] - 1$). The bottom right panel shows the faction of the total utility gain from switching to the optimal linear portfolio can be achieved with the volatility timing portfolio ($\Delta U^\sigma/\Delta U^*$) Utility gains are in wealth units (e.g. a 1% gain is equivalent to a 1% state-by-state increase in the investor life-time consumption). The utility $U[\widetilde{w}^*]$ is associated with the optimal linear portfolio $\widetilde{w}^*(\mu_t, \sigma_t^2)$. The utility $U[\widetilde{w}(\sigma_t^2)]$ is the utility associated with the optimal volatility timing portfolio $\widetilde{w}(\sigma_t^2)$, and $U[\bar{w}]$ is utility associated with the myopic buy-and-hold portfolio $\bar{w} = \overline{\mu}/(\gamma\overline{\sigma}^2)$.

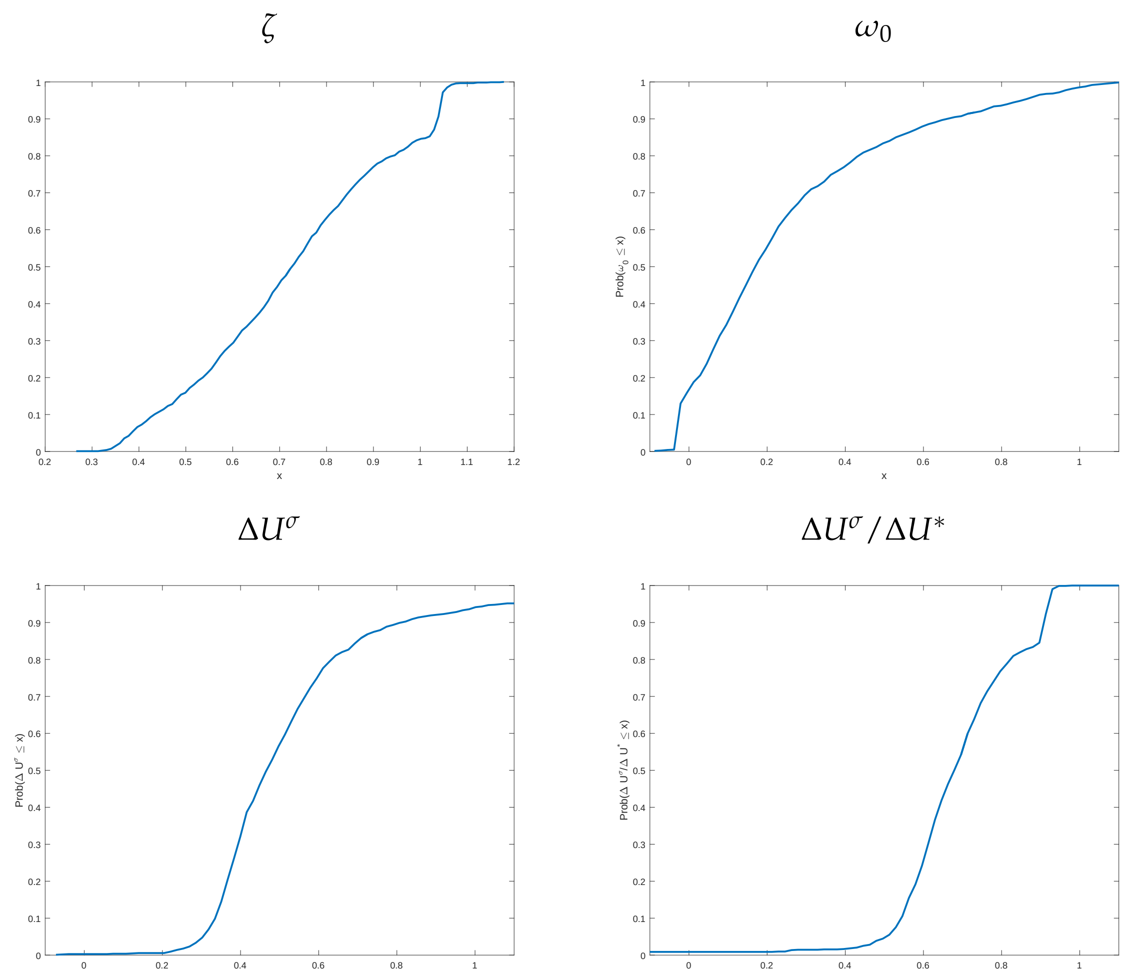
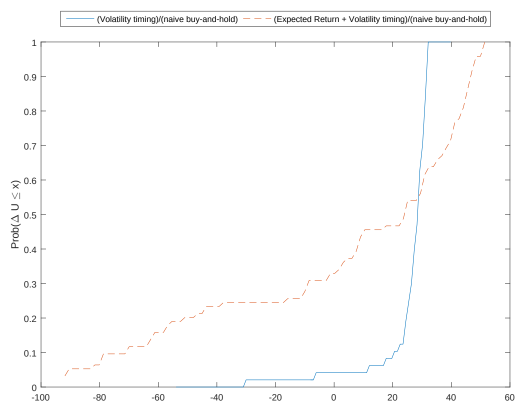


Figure 4: Distribution of utility gains due to imperfect information. This Figure shows how parameter uncertainty impacts the utility of an investor that must use a 90 year sample to estimate a forecasting model for expected returns and volatility. Section 4.2 describes the calculation in detail. The first Panel shows the cumulative distribution of utility gains relative to myopic buy-and-hold when the investor does expected return and volatility timing ($\Delta U^* = U[\widetilde{w}^*]/U[\bar{w}] - 1$) and just volatility timing ($\Delta U^\sigma = U[\widetilde{w}(\sigma_t^2)]/U[\bar{w}] - 1$). The second panel shows the cumulative distribution of $U[\widetilde{w}^*] - U[\widetilde{w}(\sigma_t^2)])/U[\widetilde{w}(\sigma_t^2)]$, the gains from switching from pure volatility timing to also timing expected return.

Gains from timing: cumulative distribution function



Gains from timing expected returns and volatility relative to only timing volatility: cumulative distribution function

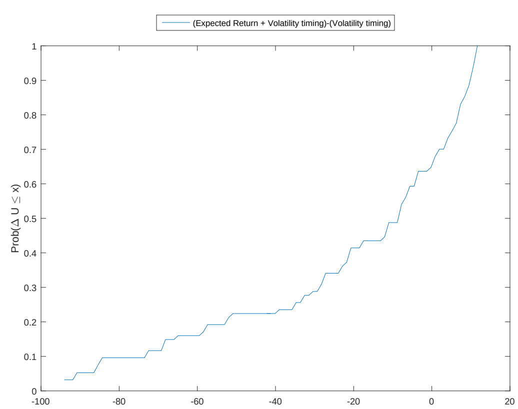
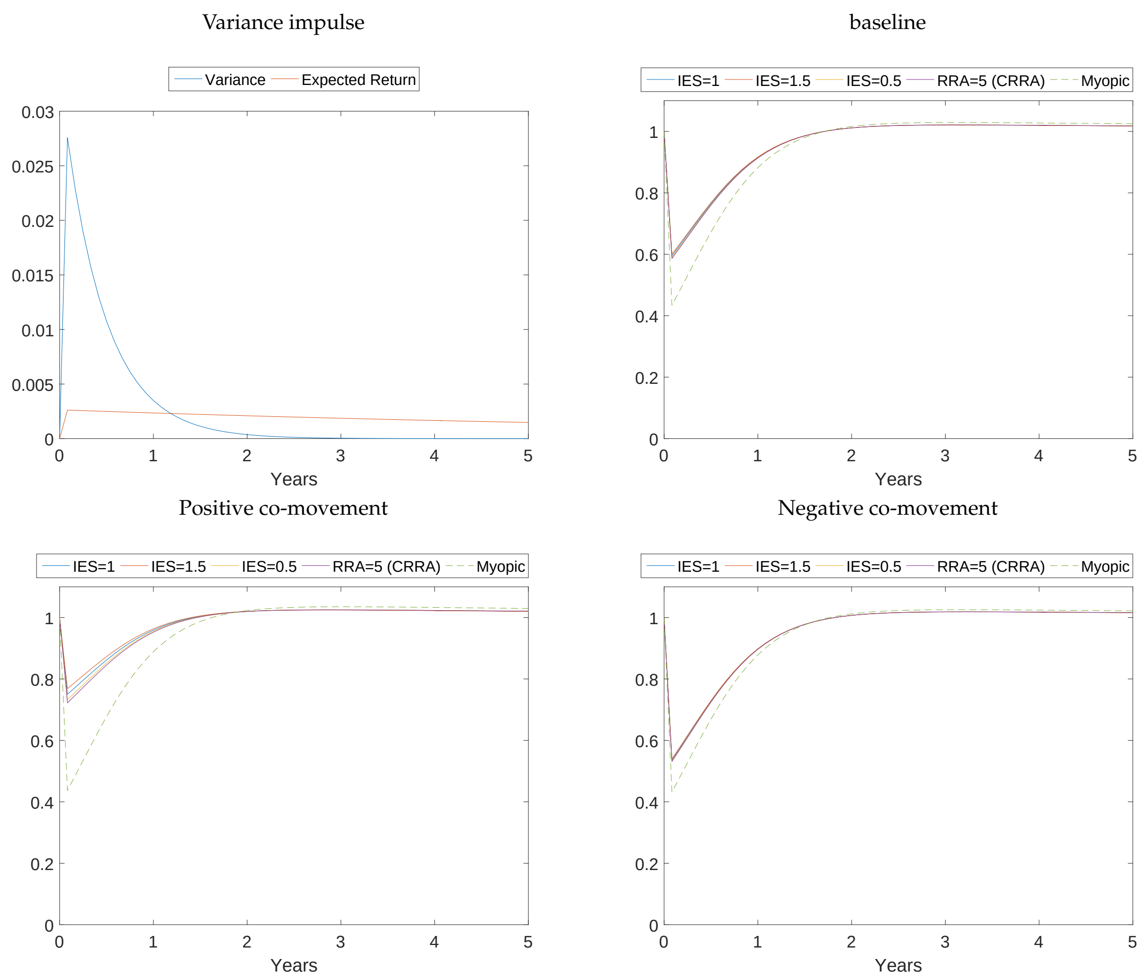


Figure 5: Alternative preference parameters. We plot the optimal response to volatility shocks as we change the EIS from 0.5 to 1 (our benchmark case) to 1.5. We also include the CARA case as well where the EIS is the inverse of relative risk aversion. We see that each choice produces quantitatively similar responses to our benchmark case of EIS=1.



. Appendix: Not intended for publication

This appendix contains additional results.

A. Numerical solution

Our solution method follows Judd (1998, Chapter 11). We first conjecture the value function $g(\mu, f(\sigma^2))$ expressed as bivariate Chebyshev polynomials of order N . We use N = 6. For our baseline parameter estimates results do not change with N up to 10, but take increasingly more time to solve. We calculate the derivatives of these functions as well as the optimal portfolio. We then plug these quantities into the HJB (12) and project the resulting residuals onto the complete set of Chebyshev polynomials up to order N. We use the built-in Matlab routine fsolve to find the coefficients of the asset price polynomials that make the projected residuals equal to zero.

B. Additional references on conventional wisdom

We include links (click for hyperlink) to additional advice on how to respond to volatility. These sources are meant to convey the conventional view given by practitioners and academics that investors should not respond to increases in volatility, and that long horizon investors may in fact want to buy stocks during periods of high volatility.

Fidelity: "A natural reaction to that fear might be to reduce or eliminate any exposure to stocks, thinking it will stem further losses and calm your fears, but that may not make sense in the long run." "Do not try to time the market." "Invest regularly despite volatility."

Charles Schwab: "Understandably, investors often become nervous when markets are volatile, and we are hearing many questions from clients." "They should also resist the urge to buy and sell based solely on recent market movements, as it could hobble their performance over time."

Forbes: "5 Tips To Survive Stock Market Volatility In Retirement."
"Stay The Course: While staying the course might sound boring to you, it is likely the absolute best thing to do right now. In fact, market volatility is the main catalyst behind a lot of bad financial behaviors – most specifically – buying high and selling low. Despite this widespread knowledge, retirees often overreact when the market drops and divest some of their equities. One way to minimize this harmful financial behavior is to hire a financial advisor. One of the great benefits of having a financial advisor is to steady your emotions during volatile markets. In fact, the retirement income certified professionals

surveyed by The American College of Financial Services stated that keeping their clients from overreacting during volatile markets was a crucial aspect of protecting their clients? retirement security.
Reduce Your Withdrawals: As a retiree, you often need to sell your investments in order to generate the income needed to meet your retirement expenses. Selling investments and taking portfolio distributions during a volatile market highlights a unique retirement risk called sequence of returns risk. Sequence of returns risk is unique because until you start withdrawing money from your investments it has no impact on your portfolio. However, when you sell stocks right after a significant market downturn, you lock in lower returns which can negatively impact the longevity of your investment portfolio. As such, if you can be flexible when markets are volatile and avoid selling as much stock, you can vastly improve how long your retirement portfolio will last. For many people, reducing market withdrawals also involves reducing expenses, even if it is just for a short period of time."

CNBC: "Investing for dummies – and smart guys – in a whipsaw market." "What may be good advice for a 21-year-old may not be the best course of action for a 65-year-old."

New York Times: "Stocks are most useful for long-term goals. So unless those goals have changed in the last few days, it probably doesn't make much sense to overhaul an investment strategy based on a blip of market activity." "Plenty of research shows that if you miss just a few days of the market's biggest gains, your long-term portfolio will suffer badly."

US News: "Volatility can also provide opportunities for investors looking for bargains. Amid an uncertain outlook for the market and key influencers such as interest rates and China, it might be a good idea to take a page from "Dr. Strangelove" – or learn to stop worrying and love the volatility."

USA Today: "Don't attempt a strategy of bailing out temporarily until things 'calm down'."