The Factor Risk in Low-Risk Anomalies*

Sina Ehsani†

November 9, 2021

Abstract

The low variance (LV) strategy always bets against the volatile leg of common factor-portfolios. The risk of the strategy, measured by factor exposures, is thus perfectly predictable based on the status of factor portfolio variances during the formation period. I find that the strategy earns alpha only when traders have to bear major factor risk to arbitrage it away. These results are consistent with models that rationalize anomalies by arbitrageurs reluctance to eliminate mispricing due to factor risk aversion. I use the framework to analyze the sources of risk and return to popular volatility trading strategies and propose a new trading strategy that uses time-series factor data to manage the variance strategy in the cross-section.

JEL Code: G11, G12, G17

Key words: Risk Return Trade-off, Volatility Effect, Variance Anomaly

*I thank Juhani Linnainmaa for helpful comments and long discussions that have greatly benefited the paper.
†College of Business, Northern Illinois University. email: sehsani@niu.edu.

1 Introduction

Academic research finds that assets within each class are priced as if the volatility premium is negative. Risk-arbitrage strategies such as low variance (LV) and betting against beta (BAB) exploit this puzzle by taking long positions in low-risk assets and short positions in riskier assets. To explain the puzzling trade-off, past studies have focused on the role of frictions and behavioral biases that, empirically, is obtained by targeting the differences in the premiums of portfolios sorted on volatility.[1] I develop a framework that recognizes the *premium* terms but also identifies important embedded *time-series* components that collectively produce the low-variance (LV) anomaly. I show that the LV strategy contains bets against the variance return trade-offs of all common risk factors. The return distribution properties stemming from these time-series factor bets are undesirable in a mean-variance sense. However, the anomaly owes its existence to these intended or unintended bets because, at times, they increase arbitrage risk by lining up with common factor loadings, preventing arbitrageurs to take the opposite position.

Because LV is a bet against factor variances, its future factor betas are predictable, highly time-varying and, every so often, substantially larger than those implied by unconditional estimates. LV's *investment* period exposure to a risk factor is determined by the variance of the long and short legs of that risk factor during the *formation* period. This two-faced behavior is inevitable because the strategy is designed to bet against (on) factors with volatile long (short) legs. Thus, even though the low variance trade has the status of a "defensive" strategy, its large factor exposures can oftentimes be in either direction. This intuition is highly consistent with

[1] Several explanations for this anomaly have been proposed: leverage constraints leave high-return-seeking investors with no choice but to bid up high beta stock prices, (Black, 1972; Frazzini and Pedersen, 2014); arbitrage risk (Pontiff, 2006) and impediments to short selling deter investors from removing mispricing (Stambaugh et al., 2015); investors may have lottery or skewness preferences (Barberis and Huang, 2008; Bali et al., 2011, 2017; Han and Kumar, 2013; Schneider et al., 2017); overconfident and unsophisticated traders invest in high beta stocks during optimistic times (Antoniou et al., 2015); institutional investors are inclined to buy high beta stocks because they are instructed to beat a benchmark (Baker et al., 2011; Christoffersen and Simutin, 2017); and divergent opinions about high beta assets makes investors prone to speculative overpricing (Hong and Sraer, 2016).

the behavior of LV with respect to the factors in Fama and French (2015): changes in Fama and French (2015) factor variances command dramatic time-variation in LV's factor betas to such an extent that betas predictably vary from -5 to 5 in just a few months.

Consider the value factor, HML, as an example to illustrate the predictability of LV factor loading. When growth stocks (the short leg of HML) are more volatile than value stocks (the long leg), LV shorts growth and invests in value, becoming a long bet on HML. Next month, value may become more volatile than growth, obligating LV to take the opposite position and short HML. Estimates of unconditional HML betas obscure the low variance strategy's conditional riskiness by averaging across positive and negative extreme factor realizations. As a result, a regression of LV on the Fama and French (2015) five-factors produces an HML coefficient of 0.07 with an insignificant t-stat of 1.44 while conditional HML betas, as was to be expected, vary from large negative to large positive and vice-versa all the time.

Not only factor betas, but future returns to the strategy also change according to the formation period status of factor variances. Expected return to the strategy is 0.60% when it is expected to bet *on* factor risks and -0.47% when it is expected to bet *against* them. This predictability in returns, however, is totally consistent with a risk-based explanation to returns because exposure to factor risk should be compensated with higher returns. What is surprising is that controlling for factor risks does not eliminate the gap: LV earns a large Fama and French (2015) five-factor alpha of 39 bps (t-stat of 3.75) when it is long factor risk and zero in other times. That is, when low variance coincides with low factor risk, the anomaly is completely arbitraged away. This finding can be best explained by the theoretical model of Kozak et al. (2018) wherein arbitragers are reluctant to trade sentiment-driven mispricings that would expose them to factor risk. LV is a prominent representation of such mispricing because it contains bets against the variance of *all* common risk factors. The returns emerging from factor bets are undesirable and bring much of the strategy's variability, yet they indirectly protect the profits stemming from other sources by distancing the strategy from a "free lunch".

Consistent with this interpretation and the empirical finding that LV earns zero alpha when factor risk is low, I find that arbitrage activity in the variance space is decreasing in factor risk, supporting the idea that factor risk may deter arbitrage. The negative relationship between arbitrage activity and factor risk implies that risk-averse rational investors who perceive the LV anomaly as an arbitrage opportunity stay away when exposure to risk factors is large. In this respect, low factor risk renders higher arbitrage activity.

I analytically describe the mechanism by which variance return trade-offs of common factors in time-series transform to the return of the cross-sectional LV strategy. I show that all assets, and in particular volatile and high beta stocks, are crucial in this transition. High beta assets amplify time-series factor trade-offs and spread them in the cross-section, producing more than 80% of the variation in LV returns. From an investment perspective, the dynamic factor exposure of the strategy impairs its desirability overall because the profits from this source come at much more variability than the profits stemming from the unconditional bet. Therefore, any technique that effectively tempers the time-series bet can improve the strategy's appeal. I show that volatility-neutralizing along the lines of Frazzini and Pedersen (2014) and volatility scaling as in Moreira and Muir (2017) tackle the dynamic exposure to aggregate variance and are extremely successful in improving LV's mean-variance efficiency. The volatility-neutralized and volatility-managed LV earn Sharpe ratios of 0.76 and 0.37, respectively, compared to 0.21 for the standard LV.

I propose a timing strategy that is as tradable as the static LV anomaly and targets timing the risk return trade-off, rather than risk or return. The timing strategy predicts LV strategy returns based on the status of the variance-return trade-offs of common risk factors in time-series. Because these trade-offs are in charge of most of the variation in price of variance in the cross-section, their forecasts are forecasts of LV returns. I find that variance return trade-offs are persistent and, thus, their current values are forecasts of expected LV returns. I develop a Risk-on Risk-off (RoRo) strategy that remains Risk-off (the LV's default status) if aggregate time-series trade-off is negative, and switches to Risk-on (buy high risk and sell low risk) otherwise. This strategy trades the same assets of the typical long-short variance strategy, but different in that it hits the risk-on/risk-off switch frequently. RoRo consistently outperforms the static LV by earning twice monthly returns at a similar standard deviation. The outperformance is highly significant and persists in recent data. I run several factor model regressions to explain the returns of the RoRo strategy. The models produce R^2s between 0% and 5%, which implies that we are capturing a totally different effect. These results show that as long as the variance return trade-offs of common factors exhibit persistence, the active strategy will be more optimal than its static counterpart. This conclusion follows logically by the definition of the RoRo strategy provided that security returns obey a factor structure.

The aggregate time-series trade-off measure is computed using the five factors of Fama and French (2015). I show that all subsets containing any combination of any number of factors predict LV returns significantly and with the correct sign. However, the measure becomes more precise in its forecasts as more factors are included, consistent with a multi-factor representation of asset returns. For example, a regression of expected LV returns on the aggregate trade-off measure constructed using one randomly selected factor produces a t-stat of 2.08. Average t-stats increase to 2.41, 2.83, and 3.23 for measures constructed using random draws of 2, 3, and 4 factors, respectively, and reaches a maximum of 3.56 for the trade-off measure that uses all five.

What explains the decline in returns to the LV anomaly over time? I highlight episodes of poor realized performance of the LV strategy. In particular, since the mid 1990s and until the end of the sample in 2018, it has earned close to zero. Presumably, this performance degradation is suggestive evidence of data snooping and post-publication drift in the context of McLean and Pontiff (2016) and Linnainmaa and Roberts (2018). I show that this underperformance is, at least, partially systematic and a result of more frequent positive realizations of the variance return trade-off of common risk factors in recent data, consistent with the notion that the performance of risky assets in the cross section is related to the pricing of systematic variance in the time series.

The rest of the paper is organized as follows. Section 2 shows that factor portfolio variances command future factor betas, arbitrage activity, return, and alpha of the low variance strategy. Section 3 derives an analytical framework to study the sources of risk and return to a variance strategy. Section 4 provides a linear factor model representation for the profits of the strategy and then provides empirical evidence to support the model's prediction. Section 5 investigates the existing techniques of trading variance in the cross-section through the lens of our model and proposes a new strategy. Sections 6 and 7 check for robustness by testing other predictions of the model. Section 8 concludes.

2 Factor risk as limits to arbitrage

Over the period of 1963 to 2018, a self-financing strategy that invests in low variance stocks and shorts high variance stocks earns an average monthly return of 0.29% with a t-stat of 1.56.[2] A regression of LV strategy returns on the five factors of Fama and French (2015) gives the following results:

$$\begin{array}{lllllll} \mathsf{LV} = 0.28 & -\,0.45\,\mathsf{MKTRF} & -\,0.47\,\mathsf{SMB} & +\,0.07\,\mathsf{HML} & +\,0.68\,\mathsf{RMW} & +\,0.58\,\mathsf{CMA} \\ t\text{-stat}: \ [2.86] & [-18.69] & [-13.70] & [1.44] & [14.15] & [8.48] \end{array}$$

The statistically significant intercept implies that the Fama and French (2015) five-factor investor can expand the unconditional mean-variance frontier of her portfolio by trading the variance strategy at reasonable factor risk, measured by unconditional factor betas. Figure 1 shows that these moderate factor beta estimates obscure the true dynamic risk of the strategy. The red line in each panel plots the time-series for the anomaly's exposure to one factor, estimated by a regression of LV returns on FF5 using a month of daily data. The same chart plots

[2] The low variance "factor" here is the intersections of 2 portfolios formed on size and 3 portfolios formed on variance, computed using one month of daily returns. The size breakpoint is the median NYSE market equity. Variance breakpoints are the 30th and 70th NYSE percentiles.

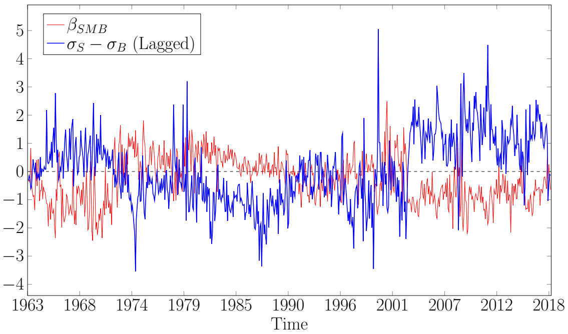
the *lagged* volatility spread for that factor, defined by last month's variance of its long leg minus that of the short leg. For example, Panel 3b of the figure displays the monthly loading of LV on the HML factor in red in addition to the *lagged* "volatility spread" between the HML's underlying portfolios in blue. The volatility spread in Panel 3b is the standard deviation of value stocks minus the standard deviation of growth stocks.[3] A positive value indicates that the value portfolio was more volatile than the growth portfolio last month. The definition for factor volatility spread here is analogous to that of factor returns; just as HML is value stock returns minus growth stock returns, HML's volatility spread is value stocks volatility minus growth stocks volatility. The strong negative relationship between HML's volatility spread and its future beta is evident in the chart: LV's exposure to HML is determined by the status of value and growth stocks volatility the month before. The same inference carries on to the remaining factor portfolios with correlations of about -0.50. Factor betas of LV during the investment period are all dual-sided with the sign and size of beta coefficients dictated by the status of the volatility spread between factor portfolios during the formation period.[4]

We formally show in later sections that this pattern in LV's exposure to factors is embedded in its returns by construction. The strategy's ex-post exposure to a factor is consistently changing in accordance to its short and long leg variances during the formation period. LV will load negatively on factors that were "long" volatility (long leg is more volatile than the short leg) during the formation period and positively on factors that were short volatility, with the magnitude of exposure increasing in factor variances. These bundle effects produce large and predictable time-varying factor exposures that impact future returns, arbitrage risk, and alpha of the strategy.

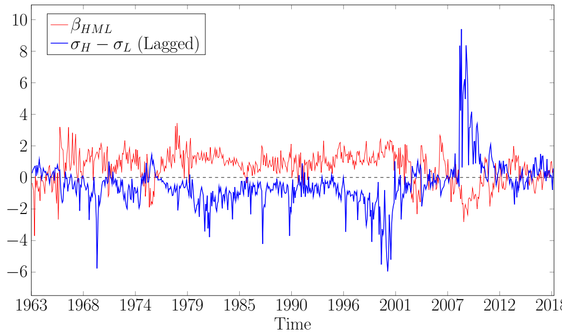
Table 1 shows the relationship between *factor variance spreads*, defined by the variance of

[3]We use standard deviation in this figure for illustration purposes. We later show that variance is the correct measure which we use in all tests.

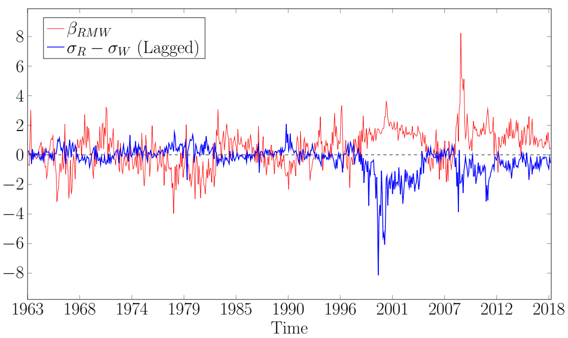
[4]For market (not presented in the figure), the long leg is always more volatile than the risk-free asset and so LV's market beta is always negative. However, the negative relationship between (ex-ante) market variance and (ex-post) LV's market beta still holds in data, implying that the strategy bets more aggressively against the market as the market factor becomes more volatile.



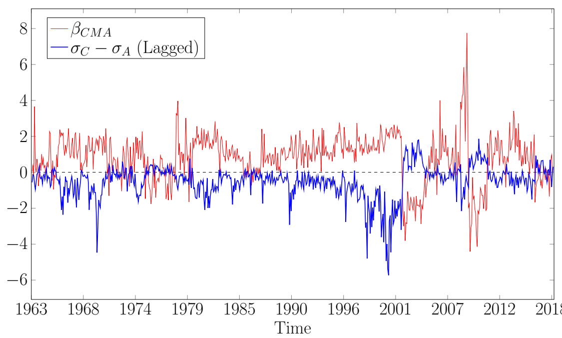
(a) SMB



(b) HML



(c) RMW



(d) CMA

Figure 1: Predictable time-varying loadings of low variance strategy.
The red line in each figure shows the time-series of monthly betas of low variance strategy with respect to each factor. The blue line shows the lagged "high-minus-low" standard deviation of each factor. The difference in the volatility of the long and short legs of a factor during the formation period predicts the exposure of the LV strategy with respect to that factor during the investment period.

a factor's long portfolio minus the variance of its low portfolio, and expected LV returns. In the first model, we regress expected LV returns on each factor's variance spread. The first row reports the result for when the independent variable is the average of all variance spreads. The estimate of -1.56 (t-stat of -3.93) indicates that one unit increase in the average factor variance spread is associated with 1.56% drop in LV returns next month. The following rows show that this result is consistent for all individual factors: LV returns can be predicted by the status of portfolio variances of all other factors; LV earns lower expected returns when the long legs of factors (small, value, robust, and conservative) are more volatile than their (large, growth, weak, and aggressive) short legs. In the last column why control for all variance spreads jointly.

All estimates remain negative and close to their univariate estimates.[5]

These results are fully consistent with a risk-based story. The LV trader is compensated for bearing positive factor risk and pays a premium when it takes on negative factor risk. Adjusting LV returns for risk, however, should eliminate the return predictability of factor variance spreads. The results in Panel B of Table 1 show that this is not the case. A one unit increase in the average variance spread between factor portfolios is associated with 0.81% lower risk-adjusted returns (t-stat of -4.06). Most individual variance spreads continue to predict LV's risk adjusted return with the correct sign. The results in Panel B indicates that the LV strategy's alpha stems from times when it aligns with factors variances, that is, when it is a bet on factor variances. A one unit increase in the average variance spread is enough to diminish more than twice of LV's alpha. These findings indicate that LV's mispricing is exclusively limited to months when the LV strategy coincides with high factor risk.

The results in Table 1 can be best interpreted by Kozak et al. (2018) who show that demand from non-rational investors produce mispricing only if arbitrageurs are exposed to factor risk when taking the other side of these trades. In their model, risk-averse rational investors trade against sentiment investors who have cross-sectional misperception about the future distribution of returns. As long as sentiment investor misperception driven demand is orthogonal to common factors, prices are not affected because arbitrageurs fully accommodate sentiment demand by taking the opposite position. Only sentiment demand that lines up with risk factors impact prices because arbitrageurs are reluctant to trade against such misperceptions aggressively.[6]

These results are also related to the model and empirical results in Herskovic et al. (2016)

[5]Because LV is a cross-sectional bet, we compute the cross-sectional variance of the market by computing the value-weighted cross-sectional variance of all stocks. The estimate is negative (-0.05) and statistically insignificant. Using the variance of the market portfolio itself also results in an insignificant estimate of 0.05. Most of our results are driven by the cross-sectional high-minus-low factors.

[6]Kozak et al. (2018) also find that the volatility anomaly correlates the most with the dominant components of asset returns. They find that the main component of asset returns describes 34% of the variation in the volatility anomaly, more than double of all the anomalies they study (Table I, page 1189). According to the model, arbitragers should be most reluctant to arbitrage the volatility anomaly.

where common variation in volatilities is a priced risk because it is positively related to the average marginal utility. However, the agents in Herskovic et al. (2016) are concerned about changes in the *level* of volatility commonality in the cross-section while the results in Table 1 show that the *direction* of the bet—on or against factor risk—is what drives the returns to the variance anomaly. These two effect are distinct. Correspondingly, Herskovic et al. (2016) find that common idiosyncratic volatility factor does not explain the volatility anomaly.

Does arbitrage activity really drop when expected factor risk is higher? We relate arbitrage activity to factor risk by associating factor betas to a measure of trade crowdedness. Lou and Polk (2012) propose a measure of arbitrage activity in momentum by computing outcome of the arbitrage process. If arbitrageurs of momentum (or any strategy) take similar positions in winner and losers stocks, their trades can produce excess return comovement in assets that are in among winners or losers. With this logic, an LV trader is more likely to simultaneously buy or sell the stocks that fall into low or high variance group. I follow Lou and Polk (2012) and measure overall arbitrage activity in the low variance strategy ($CoVol$) using the average of arbitrage activity in its long and short portfolios, where arbitrage activity in each high or low variance portfolio is the excess return correlation between all stocks in that portfolio.

Specifically, every month, stocks are sorted into tercile portfolios based on past month variance. Residuals for each stock with respect to the five factor model are calculated using a month worth of daily data. I then compute pairwise correlations between the residuals of all stocks in the high or low variance portfolio during the formation period. The average pairwise correlations are proxies of crowded trading of each variance portfolio:

$$CoVol_{LV} = \frac{1}{N} \sum_{i=1}^{N_{LV}} Corr(\epsilon_{i,LV}, \epsilon_{LV})$$

$$CoVol_{HV} = \frac{1}{N} \sum_{i=1}^{N_{HV}} Corr(\epsilon_{i,HV}, \epsilon_{HV})$$

where ϵ_i is the five factor residual of stock i, and ϵ_{HV} and ϵ_{LV} are averages of residual correlations for each portfolio, and can be interpreted as the excess correlation among high or low variance stocks. Overall arbitrage activity of the LV trade is defined by:

$$CoVol = \frac{1}{2}(CoVol_{LV} + CoVol_{HV})$$

Figure 2 shows the time-series of $CoVol$ over the sample period. The larger arbitrage activity in recent data is consistent with the exponential growth in the amount of arbitrage capital devoted to low risk traders over the past decade. According to `etfdb.com`, there are 75 low-volatility ETFs with an AUM of above \$100 billion by the end of 2019. The capital devoted to low variance strategies by mutual funds and hedge funds is even larger. Therefore, following Lou and Polk (2012), I present the results for both $CoVol$ and its detrended series that transforms $CoVol$ to a stationary series. Also, Lou and Polk (2012) use the portfolio formation period data for estimating arbitrage activity because their goal is to predict momentum peaks and reversals. I compute the measure using formation period data as in their paper and using investment period data for robustness.

Table 2 present the pairwise correlations between different $CoVol$ measures and variance spreads. A positive sign indicates that arbitrage activity increases when factor long legs are more volatile than the short legs; that is, when LV bets against factor risk and arbitraging it does not require taking on factor risk. The first column, titled "Detrended", presents the correlation between variance spreads at time t and the detrended $CoVol$ measure during time $t+1$. The correlation between the average variance spread and next month arbitrage activity is 0.23 with a t-statistics of 6.20. The following rows present pairwise correlations between individual variance spreads and detrended $CoVol$. The positive correlation between variance spreads and arbitrage activity reveals that arbitragers refrain from trading LV when it is expected to become a positive bet on factor variances.

The following columns report correlations between factor risks and other variations of $CoVol$.

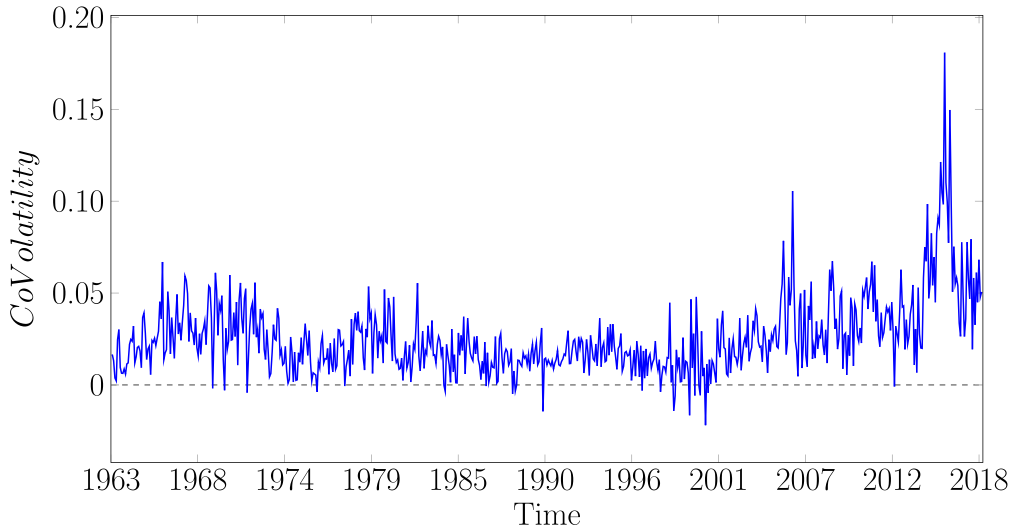


Figure 2: Arbitrage activity in low variance trade.
This figure displays the time-series of arbitrage activity ($CoVol$) as in Lou and Polk (2012). $CoVol$ is the Lou and Polk (2012) formation period measure of arbitrage activity calculated by the average of pairwise correlations between stocks in the low variance portfolio and pairwise correlations between stocks in the high variance portfolio.

The results are robust for both detrended and not detrended $CoVol$, as well as in contemporaneous tests. Most correlations are positive and the few negative are not distinguishable from zero. Taken together, the results in Table 2 are consistent with the hypothesis that arbitrage activity is decreasing in factor risk. Arbitragers trade LV when they it is not a bet on factor variances. In other times, factor risk prevents arbitragers to exert price pressure and eliminate possible mispricing.

Next, I develop an analytical framework to characterize and quantify the LV strategy's cross-sectional and time-series bets. We show that volatility spreads between factor portfolios during the formation period—that are the commanders of next month factor exposures—play a critical role in producing the risk return profile of portfolios sorted on variance.

3 An anatomy of the volatility effect

In this section we introduce a simple variance-based strategy and decompose its profits. The purpose of this exercise is three-fold. First, we illustrate that any variance-based strategy contains two inherent bets against variance, one in the cross-section and one in the time-series. We then explore the properties of each component and show that the time-series component generates substantial volatility but little profits. Finally, we show that the time-series component is entirely systematic and derives from LV's time-varying exposure to factor risk.

3.1 A risk-based trading strategy

Consider a strategy that uses the information contained in past variance for security selection. At time t, weight assigned to asset p is calculated using the identity[7]

$$w_{p,t} = \left(\mathsf{var}_{p,t-1} - \overline{\mathsf{var}_{t-1}}\right) \tag{1}$$

where $\mathsf{var}_{p,t-1}$ is the variance of asset p at time $t-1$, computed using a month of daily returns, and $\overline{\mathsf{var}_{t-1}}$ is the time $t-1$ average variance of all assets in the market. The overbar represents the cross-sectional average throughout the paper. The return of the strategy at time t is given by[8]

$$V^p = \left(\mathsf{var}_{p,t-1} - \overline{\mathsf{var}_{t-1}}\right) r_{p,t} \tag{2}$$

where $r_{p,t}$ is the realized return of portfolio p at time t. This simple strategy holds assets in proportion to their variance relative to the average variance of all assets in the market such that if implemented in a cross-section of P portfolios, it creates a zero-cost investment. The strategy invests heavily in portfolios with high past volatility and finances this position by shortselling portfolios with lower than average volatility. The relevance of the strategy relies on the premise that an asset's variance relative to the average includes useful information about its future performance. Assuming test portfolios are exposed to systematic risk (i.e., p is a well-diversified

[7]To be more precise, we use $100 \times w_{p,t}$ for the weight of decile portfolios, and $10 \times w_{p,t}$ for the weight of percentile portfolios. Because weights can be arbitrarily scaled to obtain any premium, we mostly rely on the contribution of each component relative to the total effect in our analysis.

[8]The trading rule in (2) is inspired by studies such as Lo and MacKinlay (1990), Lehmann (1990), and Conrad and Kaul (1998) who use the linear weight rule to study the profits to contrarian trading strategies. The trading rule in (2) represents a long variance strategy which can be modified to a short variance strategy by adding a negative sign preceding the expression.

portfolio), it is not surprising or against the random walk model for this expression to earn positive profits: if variance is a good proxy for future risk, and risk commands differences in expected returns, the strategy should outperform zero. Contrarily, a negative premium is diametrically opposed to market efficiency and can be translated into abnormal risk-adjusted profits.

Previous studies such as Ang et al. (2006, 2009) use an alternative approach that forms a self-financing low-minus-high strategy using two extreme portfolios sorted on variance. In this section we depart from past studies that focus on the spread between the top and bottom portfolios in two ways. First, our strategy trades the entire cross-section, not just the extreme portfolios. Second, although the strategy is a zero-cost investment, portfolio weights change every month according to their realized variance, whereas the "low minus high" approach always invests a unit in Portfolio 1 financed by a short position of the same size in Portfolio 10. Although highly correlated, our variant definition of variance profits expressed in equation 2 has analytical advantages that permits us to quantify all potential sources of profits. We thus employ the linear weighting schedule to learn and draw inferences about the size and origins for each source of profits. However, the following sections, when it comes to trading implications, we explain and predict the returns of the standard HML-type LV factor.

3.2 Analytical decomposition of profits

Denote by μ_p and σ_p^2 the unconditional expected return and variance of portfolio p. Averaging the profits of the trading program of equation 2 in the cross-section of P portfolios and taking expectation gives:

$$\begin{aligned} \mathbb{E}[V^{CS}] &= \mathbb{E}\big[\frac{1}{P}\sum_{s=1}^{P}\left(\mathsf{var}_{p,t-1} - \overline{\mathsf{var}_{t-1}}\right) r_{p,t}\big] \\ &= \frac{1}{P}\sum_{p=1}^{P}\mathsf{Cov}\left(\mathsf{var}_{p,t-1}, r_{p,t}\right) - \mathsf{Cov}\left(\overline{\mathsf{var}_{t-1}}, \overline{r_t}\right) + \frac{1}{P}\sum_{p=1}^{S}\left(\sigma_p^2 - \overline{\sigma^2}\right)\left(\mu_p - \overline{\mu}\right) \end{aligned} \tag{3}$$

Equation 3 can be expressed in matrix notation to separate three sources of profits (or losses):

$$\mathbb{E}[V^{CS}] = \frac{1}{P}\mathsf{Tr}(\Omega) - \frac{1}{P^2}1'\Omega 1 + \sigma_\mu^2$$
$$= \underbrace{\frac{P-1}{P^2}\mathsf{Tr}(\Omega)}_{\text{Own-covariance}} - \underbrace{\frac{1}{P^2}(1'\Omega 1 - \mathsf{Tr}(\Omega))}_{\text{Cross-covariance}} + \underbrace{\frac{1}{P}\sum_{p=1}^{P}\left(\sigma_p^2 - \overline{\sigma^2}\right)\left(\mu_p - \overline{\mu}\right)}_{\text{Premium}} \quad (4)$$

where $\Omega = \mathbb{E}\left[(\mathsf{var}_{p,t-1} - \overline{\sigma^2})(r_{p,t} - \overline{\mu})'\right]$ is the 1st order cross-covariance matrix of variance and returns:

$$\Omega = \begin{bmatrix} \mathbb{E}\left[(\mathsf{var}_{1,t-1} - \overline{\sigma^2})(r_{1,t} - \overline{\mu})\right] & \mathbb{E}\left[(\mathsf{var}_{1,t-1} - \overline{\sigma^2})(r_{2,t} - \overline{\mu})\right] & \dots & \mathbb{E}\left[(\mathsf{var}_{1,t-1} - \overline{\sigma^2})(r_{n,t} - \overline{\mu})\right] \\ \mathbb{E}\left[(\mathsf{var}_{2,t-1} - \overline{\sigma^2})(r_{1,t} - \overline{\mu})\right] & \mathbb{E}\left[(\mathsf{var}_{2,t-1} - \overline{\sigma^2})(r_{2,t} - \overline{\mu})\right] & \dots & \mathbb{E}\left[(\mathsf{var}_{2,t-1} - \overline{\sigma^2})(r_{n,t} - \overline{\mu})\right] \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{E}\left[(\mathsf{var}_{P,t-1} - \overline{\sigma^2})(r_{1,t} - \overline{\mu})\right] & \mathbb{E}\left[(\mathsf{var}_{P,t-1} - \overline{\sigma^2})(r_{2,t} - \overline{\mu})\right] & \dots & \mathbb{E}\left[(\mathsf{var}_{P,t-1} - \overline{\sigma^2})(r_{P,t} - \overline{\mu})\right] \end{bmatrix}$$

$\mathsf{Tr}(.)$ is the trace of matrix, and $\frac{1}{P}\sum_{p=1}^{P}\left(\sigma_p^2 - \overline{\sigma^2}\right)\left(\mu_p - \overline{\mu}\right)$ captures the cross-sectional dispersion in unconditional returns of portfolios sorted on past variance.

According to equation 4, the negative volatility effect can emerge because of one or more of the following effects:

1. Negative own-covariance: A portfolio's high volatility signals lower future returns
2. Positive cross-covariances: A portfolio's high volatility signals future higher returns on other portfolios
3. Differences in mean returns: the persistent difference in returns across variance sorted portfolios

The third term is independent of the time series dynamics between variance and returns. Existing research that uses frictions and constraints as potential explanations to the LV anomaly, implicitly, studies the third term of the decomposition.

3.3 Estimating the profit decomposition matrix

We implement the trading program in equation 2 on decile portfolios sorted on variance and derive the decomposition matrix of equation 4. Appendix A.2 provides a description of data and our test assets. Table 3 shows the average profits of the linear weighting strategy and its components. Panel A shows the aggregate values and Panel B presents the portfolio level contributions. The column labeled "Strategy" shows that over the period of 1963 to 2018, the strategy earns an average premium of -2.93% with a t-stat of -1.19. An inspection of portfolio level contributions in Panel B reveals that most portfolios contribute to the negative price of volatility (8 out of 10 have negative signs). Only portfolios 6 and 7 are priced as if the price of variance is positive, implying that the negative volatility effect is a common feature of the cross-section. This observation is noteworthy considering that most papers study the volatility effect by focusing on the two extreme portfolios.

The first component of equation 4 that can generate the negative variance effect is titled "Own". This term captures the impact of own-covariance between past variance and future returns. Panel A shows that this term has a large total effect of -4.49%. When we breakdown the "Own" contribution by each portfolio, two important results emerge. First, variance is negatively correlated with future returns of decile portfolios without exception, meaning that across all variance sorted portfolios, a large volatility signals a lower return next month. Second, the size of own-covariance increases dramatically from -0.03% for portfolio 1 to -1.98% to portfolio 10. The aggregate contribution of -4.49%, however, lacks statistical significance.

The second component of equation 4 captures the cross-covariance impact, which is the covariance between the variance of a portfolio and expected returns of other portfolios. The trader that buys high variance by shorting other assets will benefit if high variance of that portfolio is followed by lower returns in other assets. The results in column titled "$-1\times$Cross" show that this is the case, high variance of a portfolio is followed by lower returns on other portfolios, reflected in the negative aggregate cross-covariance (we present $-1\times$ the value). Portfolio level contributions show that high variance of different portfolios are followed by lower returns on other portfolios without exception. As a result, cross-covariances attenuate the negative volatility effect and, in total, it offsets 3.24% of the negative volatility effect. Further, unlike own-covariances that were strongest among high variance portfolios, such asymmetry is absent here. Instead, cross-covariance impacts are somewhat uniform, meaning that past variance of a portfolio covaries with future returns of other portfolios similarly. This discrepancy between the impact of own- and cross-covariances across deciles explains the overall larger contribution of the volatile portfolios compared to the rest. Together, the time-series components produce 42% (the difference between $+153\%$ and -111%) of the negative volatility premium.

The last component of equation 4 is titled "Premium" and captures the unconditional differences in mean returns. This term can be interpreted as the profits to a pure high variance minus low variance strategy that has hedged the time-series component. If variance of individual stocks is a good proxy for their riskiness, and if risk commands expected returns, then filtering the time series terms would curtail the "noise" and leave us with a cleaner *positive* association between risk and return. Theory predicts that high variance portfolios would earn higher returns compared to low variance portfolio in absence of any time series predictability. Contradictory, we find that the unconditional premium is *negative* and significant (-1.68% with a t-stat of -3.70). This means that removing the time-series components deepens the puzzle significantly. The pattern in portfolio level contributions to the premium components is quite different from that of the time-series components. The low variance portfolios are the predominant source for generating the unconditional profits of the variance anomaly.[9]

To summarize, we find that the negative variance effect is a pervasive property of the cross-section. The volatility effect can be captured by trading almost any segment of the market but the channel through which each portfolio contributes to the effect is different. The large

[9]We estimate the "Premium" component under the assumption that mean returns and mean variance of each portfolio is constant over our sample period. Thus, μ_p and σ_p^2 are the average monthly return and the average monthly variance of portfolio p over the entire period.

contribution of volatile stocks to the puzzle is through the time-series channel. At the opposite end of the volatility spectrum are the low variance portfolios who earn a high unconditional mean and contribute to the anomaly through the premium channel. Next, we show that all time-series effects—own and cross-covariances—are systematic, stemming from time-varying exposure to factor risk.

3.4 Breaking down the components into systematic and residual

We distinguish between the impacts of systematic and residual variances on the expected profits of the total variance strategy. This allows us to classify each premium or time-series component into systematic (related to the beta arbitrage class of anomalies) and idiosyncratic (the idiosyncratic volatility puzzle) risks. Letting s and ϵ denote the systematic and idiosyncratic subscripts, we can rewrite the *variance–return* covariance matrix in *systematic and residual variance–return* covariance matrices,

$$\begin{aligned}
\mathbb{E}[V^{CS}] =& \\
=& \underbrace{\frac{P-1}{P^2}\mathsf{Tr}(\Omega_s)}_{\substack{\text{Own-covariance}\\ \textit{Systematic}}} - \underbrace{\frac{1}{P^2}\left(1'\Omega_s 1 - \mathsf{Tr}(\Omega_s)\right)}_{\substack{\text{Cross-covariance}\\ \textit{Systematic}}} + \underbrace{\frac{1}{P}\sum_{p=1}^{P}\left(\sigma^2_{p,s} - \overline{\sigma^2}\right)\left(\mu_p - \overline{\mu}\right)}_{\substack{\text{Premium}\\ \textit{Systematic}}} \\
&+ \underbrace{\frac{P-1}{P^2}\mathsf{Tr}(\Omega_\epsilon)}_{\substack{\text{Own-covariance}\\ \textit{Residual}}} - \underbrace{\frac{1}{P^2}\left(1'\Omega_\epsilon 1 - \mathsf{Tr}(\Omega_\epsilon)\right)}_{\substack{\text{Cross-covariance}\\ \textit{Residual}}} + \underbrace{\frac{1}{P}\sum_{p=1}^{P}\left(\sigma^2_{p,\epsilon} - \overline{\sigma^2}\right)\left(\mu_p - \overline{\mu}\right)}_{\substack{\text{Premium}\\ \textit{Residual}}} \qquad (5)
\end{aligned}$$

To estimate this matrix, we need to make an assumption regarding the volatility generating process of asset returns. For the empirical tests of this section, we assume that exposure to the five factors sufficiently describes the cross-sectional differences in systematic volatility. This is a reasonable assumption because the model explains between 76% and 95% of the variation in excess returns of variance sorted portfolios. For each decile portfolio, I estimate pre-ranking betas every month by running this regression:[10]

$$r - r_f = \alpha + \beta_m \,\mathsf{MKTRF} + \beta_s \,\mathsf{SMB} + \beta_h \,\mathsf{HML} + \beta_r \,\mathsf{RMW} + \beta_c \,\mathsf{CMA} + \varepsilon$$

Systematic (var_s) and idiosyncratic (var_ϵ) variances for each portfolio are then calculated by the variance of the estimates for the fitted values and residual terms, multiplied by the number of trading days in month.

Table 4 gives a detailed account of variance profits broken down by systematic and idiosyncratic risks according to equation 5. First and second columns show estimates for the covariance between past systematic and future returns and past residual variance and future returns, respectively. The overall and portfolio level impacts of the residual components are very close to zero. The own-covariances deriving from systematic variance are comparable to our previous estimates of own-covariances produced by total variance in Table 3. Estimates for cross-covariances follow the same pattern, all time-series effects derive from systematic variance.

Comparing the statistics of the systematic terms in this table to those of Table 3, we conclude that systematic risk governs the time series dynamics between past variance and future returns. It is in charge of the contribution asymmetry between high variance and low-variance own-covariances, and the evenly distributed contribution of cross-covariances. Residual variance does not generate any notable own- or cross-covariance effect. In summary, we find that all time-series patterns are produced by systematic variance, paving the way to link the returns of the cross-sectional low volatility strategy to the underlying price process.[11] We next derive a factor model representation of variance profits that ties our findings together.

[10]There are numerous methods for estimating portfolio or individual stock betas. Some use lower frequency returns and longer windows (Fama and French, 1992), while others such as Ang et al. (2006) use one month of daily data. I run the regressions using a month worth of daily returns to be consistent with the one-month formation period of our strategy and record the beta coefficients and residuals.

[11]Aggregating many stocks into decile portfolios eliminates most of the variation in the firm-specific component of returns. Perhaps our portfolios have very little residual variance to begin with. To alleviate this concern, we create a larger cross-section of 100 variance sorted portfolios and repeat the decomposition exercise. Appendix A.3 shows that those results are almost identical to the results presented here.

4 Factor model representation of variance profits

This section derives a factor model representation of variance profits. This should be viewed as a descriptive model that provides a concrete framework for finding the underlying source of the main result described in the last section that all time-series relationships between expected return and variance stem from systematic variance. The factor model representation also sheds light on the motivation behind the existing proposed methods for trading variance in the cross-section.

Once again, assume that the investment in each stock is equal to the *negative* of its variance relative to the average,

$$LV_t = -\big(\mathsf{var}(r_{s,t-1}) - \overline{\mathsf{var}(r_{t-1})}\big)\ r_{s,t} \tag{6}$$

I start with a single factor process for stock returns because the interpretations will become particularly straightforward in a CAPM-like setting:

$$r_{s,t} = \beta_s\, r_{M,t} + \varepsilon_{s,t} \tag{7}$$

where r_s and r_M are excess returns on the stock and the market, respectively, and $\varepsilon_{s,t}$ is the stock specific component that should carry no risk premium in absence of arbitrage. Because exposure to market risk captures the cross-section of expected returns, the unconditional expected excess return of stock s becomes

$$\mu_s = \beta_s\, \mathbb{E}[r_{M,t}]$$

The residual and market components are orthogonal to one another and sum up to total variance:

$$\mathsf{var}(r_{s,t-1}) = \beta_s^2\, \mathsf{var}(r_{M,t-1}) + \mathsf{var}(\varepsilon_{s,t-1})$$

Given our earlier results that residual variance of a portfolio is not associated with future returns of that or other portfolios in time-series, we can reasonably assume:

$$\mathsf{Cov}\left[\mathsf{var}(\varepsilon_{k,t-1}), \varepsilon_{l,t}\right] = 0, \qquad \forall\, k, l \tag{a}$$

$$\mathsf{Cov}\left[\mathsf{var}(\varepsilon_{k,t-1}), r_{f,t}\right] = 0, \qquad \forall\, k, f \tag{b}$$

$$\mathsf{Cov}\left[\mathsf{var}(r_{f,t-1}), \varepsilon_{k,t}\right] = 0, \qquad \forall\, k, f \tag{c}$$

Applying the return process of equation 7 to the linear strategy of equation 6 and taking expectations gives

$$\begin{aligned}\mathbb{E}[LV_t] &= \mathbb{E}\left[-\left(\mathsf{var}(r_{s,t-1}) - \overline{\mathsf{var}(r_{t-1})}\right) r_{s,t}\right] = \\ &= -(\beta_s^2 - \overline{\beta^2})\,\mathsf{var}(r_M)\,\mu_s - \left(\mathsf{var}(\varepsilon_s) - \overline{\mathsf{var}(\varepsilon_s)}\right)\mu_s - (\beta_s - \overline{\beta})(\beta_s^2 - \overline{\beta^2})\,\mathsf{Cov}\left(r_{M,t}, \mathsf{var}(r_{M,t-1})\right)\end{aligned} \tag{8}$$

Assuming market beta of individual stocks is normally distributed $\left(\beta_{s,f} \sim \mathcal{N}(\overline{\beta}_f, \sigma^2_{\beta_f})\right)$, the profits of the variance strategy in the cross-section of N ($N \to \infty$) stocks equal:

$$\begin{aligned}&\mathbb{E}[LV_t^{CS}] = \\ &\underbrace{-\frac{1}{N}\sum_{s=1}^{N}(\beta_s^2 - \overline{\beta^2})\,\mathsf{var}(r_M)\,\mu_s}_{\text{Pure beta arbitrage}} - \underbrace{\frac{1}{N}\sum_{s=1}^{N}\left(\mathsf{var}(\varepsilon_s) - \overline{\mathsf{var}(\varepsilon_s)}\right)\mu_s}_{\text{Pure CAPM idiosyncratic risk arbitrage}} - 2\,\overline{\beta}\,\sigma^2_{\beta}\underbrace{\mathsf{Cov}\left(r_{M,t}, \mathsf{var}(r_{M,t-1})\right)}_{\text{Market's risk return trade-off}}\end{aligned} \tag{9}$$

The first two terms capture dispersion in unconditional mean returns arising from "pure" beta and residual variance puzzles. In earlier decompositions we found that both of these components contribute significantly to the variance anomaly.

The last term reveals the mechanism by which the risk/return covariance structure of the market portfolio leads to the variance premium. This component succinctly aggregates the own-covariance and cross-covariance components deriving from systematic variance and puts it in a factor model representation: a cross-sectional strategy based on past volatility is an indirect bet against the covariance between market's variance and expected return in time-series. Asset betas transmit this bet to the cross-section. We cautiously term the covariance between market variance and expected returns "the intertemporal risk return trade-off".[12]

The intuition is straightforward. Variance dispersion across variance sorted portfolios is increasing in market variance. When market variance is high, the low-minus-high variance strategy becomes a more aggressive bet against market variance. If the covariance between market variance and expected returns is negative, market returns will be less than average the following month, and these lower returns head into the cross-section, translating into the underperformance of risky assets. Therefore, the negative covariance between market's variance and expected return is a potential source of profits to a cross-sectional low-variance strategy.[13]

It is plausible to assume that factor variances do not covary with future returns of other factors. We can extend the expression to accommodate a multi-factor return process:

$$\mathbb{E}[V_t^{CS}] = \\ = -\frac{1}{N}\sum_{s=1}^{N}\sum_{f=1}^{F}(\beta_{s,f}^2 - \overline{\beta^2})\,\mathsf{var}(r_f)\mu_s - \frac{1}{N}\sum_{s=1}^{N}\big(\mathsf{var}(\varepsilon_s) - \overline{\mathsf{var}(\varepsilon_s)}\big)\mu_s - 2\sum_{f=1}^{F}\overline{\beta}_f\sigma_{\beta_f}^2\,\mathsf{Cov}\big(\mathsf{var}(r_{f,t-1}), r_{f,t}\big) \tag{10}$$

Equation 10 provides a rich framework to study the LV strategy. According to this equation, when returns obey a multi-factor structure, return to an LV strategy is determined by the status of factor risk return trade-offs, plus two unconditional bets on residual and systematic variance.

[12] The terminology can be motivated using the vast literature that examines the relationship between conditional variance and expected returns, that mostly focus on the market factor. I provide a brief review of this literature in appendix A.4.

[13] The premium emerging from time-series component was not statistically significant in profit decompositions of the variance strategy. The absence of statistical significance stems from the highly time-varying covariance between market variance and expected returns, a property that is in accord with prior studies that document the counter-cyclical variation of the intertemporal risk return trade-off.

We found in profit decompositions that the unconditional bets produce profits and time-series factor bets produce variability. We next study the well-known techniques of exploiting low risk anomalies within this framework.

5 Trading variance in the cross-section

The linear factor model representation of LV returns and the empirical results from profit decomposition put forward a framework to study different techniques for trading variance in the cross-section. I first show that these techniques try to deal with the undesirable time-series component in one way or another and then propose a strategy of timing LV based on timing factors' risk-return trade-offs.

5.1 Volatility neutralizing

I adopt the approach of Frazzini and Pedersen (2014) to study the impact of volatility-neutralizing on LV returns. Frazzini and Pedersen (2014) propose a beta-neutral factor to capture the beta effect. The standard approach of capturing the beta effect or any anomaly in the literature is to form a self-financing portfolio that buys a unit of a portfolio financed by selling a unit of another. Instead, Frazzini and Pedersen (2014) first finance the high and low beta portfolios with the risk-free asset, and then lever or de-lever each portfolio using its past beta such that the ex-post realized beta of each portfolio equal one. They then create a long-short portfolio that goes long the new low beta and shorts the new high beta portfolio. Frazzini and Pedersen (2014) find that the resulting BAB portfolio is beta-neutral, yet it earns a very large premium.

Our framework gives an explanation to the astonishing performance of BAB: by hedging out time-series exposure to market risk, BAB removes a part of the embedded time-series bet from the low-beta strategy and becomes a cleaner bet on unconditional components that are in charge of the strategy's alpha. In the spirit of the beta neutral BAB's, I use the following weighting schedule for each self-financed leg to obtain volatility-neutrality for the LV strategy:

$$\begin{aligned} w_{L,t}^{VN} &= \frac{1}{\sqrt{D_{t-1}} \times SD_{L,t-1}} \\ w_{H,t}^{VN} &= \frac{1}{\sqrt{D_{t-1}} \times SD_{H,t-1}} \end{aligned} \tag{11}$$

where w_t^{VN} is the weight of the portfolio designed to produce a volatility-neutral long-short portfolio at time t, and D and SD are the number of trading days and standard deviation of returns, respectively.[14] I then finance each leg with the risk-free asset and lever or de-lever each self-financed portfolio using its last month variance. The return to this strategy is given by:

$$r_t^{VN} = w_{L,t}^{VN}(r_L - r_f) - w_{H,t}^{VN}(r_H - r_f) \tag{12}$$

The idea is to separate the unconditional component of the strategy from exposure to aggregate variance. If past variance is a good predictor of future variance, then this weighting program should produce a portfolio that is neutral to time-series changes in aggregate volatility. The difference between the standard LV and its volatility-neutral version is the unintended time-series component: $w_{L,t}^{TS} = 1 - w_{L,t}^{VN}$ and $w_{H,t}^{TS} = 1 - w_{H,t}^{VN}$. Return to the time-series component, that we denote by r^{TS} will be:

$$r_t^{TS} = w_{L,t}^{ST}(r_L - r_f) - w_{H,t}^{TS}(r_H - r_f) \tag{13}$$

By construction, the summation of return to the vol-neutral and time-series components add up to the return on the standard low-minus-high strategy:

$$r_t^{LMH} = r_t^{VN} + r_t^{TS}$$

[14] The low and high variance legs here are HML-type portfolios: the long leg is the average of 3 small-low variance, mid-low variance, large-low variance, and the short leg is the average of 3 small, mid, and large high variance portfolios.

Table 5 shows the breakdown of LV returns by its volatility-neutral and time-series components. We see that the technique extracts the unconditional component from LV returns effectively. The mean returns and associated t-stats are consistent with the "premium" and "time-series" estimates derived from earlier decomposition. By hedging out the time-series bets, we are able to retain more than 2/3 of the premium and reduce standard deviation by more than 80 percent. As a result, the strategy now earns an impressive t-stat of 5.71. Further, the much larger t-stat of the vol-neutral LV compared to the standard LV should be accompanied by little factor risk because all the factor risk is hedged out and transferred to the TS component.

Our results prove that hedging out the time-series bet on variance—which takes the form of factor risk—is the root of the remarkable performance of our vol-neutral LV and Frazzini and Pedersen's BAB. The volatility-neutral LV strategy, just as it is the BAB factor, is like a near arbitrage opportunity on paper. While it sounds appealing at first sight, Novy-Marx and Velikov (2019) show that implementing vol-neutral LV in practice is problematic due to its use of leverage, risk-free rate financing, and large positions in illiquid and small stocks.

5.2 Volatility scaling

Another proposed approach for improving mean-variance efficiency of factor portfolios is volatility-management as in Fleming et al. (2001), Moreira and Muir (2017), and Barroso and Maio (2018). These papers use the absence of risk-return trade-off in time-series of common risk factors to motivate a volatility managing strategy. Fleming et al. (2001) show that we can form volatility timing strategies that take less risk when volatility is higher to outperform unconditionally efficient static portfolios. Moreira and Muir (2017) extends this result to a set of nine major risk factors. The idea is that because variances can be estimated with greater precision than expected returns, we can assume expected returns are constant and let the time-variation in risk be the only determinant of the time-variation in portfolio weights. I confirm the main findings of these papers for the LV strategy. Using the opposite of past month's variance as the weight assigned to next month increases the Sharpe ratio of the standard low-variance factor from 0.21 to 0.37. As a result, a regression of vol-managed LV on five-factors produces a larger and more significant intercept with a t-stat of 5.12 compared to 2.86 of the original LV.

The success of volatility managing LV can be systematically justified in our framework. Vol-managed LV enhances the unconditional bet by tempering the aggressive time-series bets against factor variances. The resulting portfolio becomes less related to factors and closer to its "pure alpha" unconditional component. Consistent with this explanation, FF5 explains only 35% of our vol-managed LV but 75% of the standard LV. Further, the vol-managed version correlates less with the time-series component of LV strategy ($\rho = 0.56$ vs 0.98 for standard LV) and more with the unconditional component ($\rho = 0.19$ vs 0.16 for standard LV).

Despite the intuitive motivation for volatility-managing portfolios, in their comprehensive study, Cederburg et al. (2019) show empirically that volatility-scaling long-short portfolios can not be distinguished from scaling portfolios by a random variable. They find that the number of anomalies that benefit or deteriorate from volatility management is split in half. My findings above justify volatility-management for the class of anomalies that use return variation as the sort variable. For this group, volatility-managed versions should systematically outperform their corresponding unmanaged counterparts in a mean-variance sense. In fact, Cederburg et al. (2019) show improvements in Sharpe ratios of the anomalies constructed on past market beta (two anomalies, $\beta - D$ and $\beta - FP$) and return volatility (two anomalies, $Ivol$ and $Tvol$).[15]

5.3 A new signal for timing LV

My benchmark model is static and so the statistics we used for estimating the components of the variance effect relied on the total sample. In a more general form, wherein all terms are time-varying, time-series variation in trade-off should be reflected in variation in LV returns in real-time. Should this be the case, forecasts of factor trade-offs in time-series would also be

[15] In Cederburg et al. (2019), $Svol$ is labeled systematic volatility and is not among the above list. However, their $Svol$ is not a sort on variance but a sort on VIX betas and, therefore, not subject to the factor variances.

forecasts of returns to the cross-sectional LV strategy.

5.3.1 Predicting common factor trade-offs

Variance return trade-off of common factors are hard to measure, let alone prediction. A large literature—mostly focused on market portfolio—attempts to just estimate the static or time-varying trade-offs by assuming alternative specifications for the mean, the volatility process, or the intertemporal hedging demands. We use a simple proxy for this tradeoff: the product of variance and return of common factors. This is an obvious choice considering that covariance between variance and returns is the last component of equation 10. Specifically, I compute time t aggregate common factor variance return trade-off by:

$$\mathsf{Tradeoff_t} = -1 \times \frac{\sum_f \left[\mathsf{var}(r^f_{H,t}) - \mathsf{var}(r^f_{L,t})\right] \times r^f_t}{\sum_f \left|\mathsf{var}(r^f_{H,t}) - \mathsf{var}(r^f_{L,t})\right|} \tag{14}$$

where f denotes a factor. We use the variance of a factor's high portfolio minus the variance of the factor's low portfolio because it is the covariance between *variance of high minus variance of low* and *high-minus-low returns* that takes the form of LV returns, and therefore the proper measure.[16] We further divide the product of risk and return by total absolute values of factor variances for ease of interpretations but the results presented in this section are entirely driven by the numerator. The measure is simple to interpret: a high variance followed by a positive return is a positive realization of variance return trade-off, which should be concurrent with lower returns to the LV strategies, and vice versa. Also, the purpose of scaling by -1 is to have a measure that positively correlates with the LV strategy—which is a bet against factor trade-offs. Therefore, our Tradeoff measures the negative of common factor variance return trade-offs.

Our predictor of LV returns is simple to calculate, can be computed for any risk factor because it only uses past return data, and does not impose any restriction on the shape of the

[16] Another way to obtain the measure is to adjust factor returns such that they are always "pro-cyclical" and then compute the variance return trade-off of the resulting factor. In that case, variance becomes $(\mathsf{var}(r_{H,t} - r_{L,t}))$, but factor return will be adjusted in accordance to the factor's last month variance spread ($r^* = sign\left[\mathsf{var}(r_{H,t}) - \mathsf{var}(r_{L,t})\right] \times r^f_t$). The variance-return product of this new measure is same as the one in equation 14.

risk-return relation. This informal measure selection approach is not designed to find the best possible predictor of LV returns, but intends to compute the closest match to the last component of equation 10. We do not argue that the measure is, in any sense, the only or the best proxy. Any methodology that estimates a time-varying intertemporal variance return trade-off of factor returns can be used in the following tests.

The constituents of Tradeoff, factor mean and variance, have persistence properties that enable us to test the model's prediction by instrumenting lagged values of Tradeoff for its future values. Engle (1982) and numerous follow-up studies develop one-period forecasts of aggregate variance using the autocorrelation structure of volatility. The assumption underlying ARCH and GARCH models is that volatility has a tendency to stay in its current high or low regime. Likewise, factor means exhibit weak serial correlation (Arnott et al., 2018; Ehsani and Linnainmaa, 2021). If the persistence in mean and variance perpetuate their covariance, past covariance between variance and return is well suited to predict its future realization. We find this to hold in data with a first order autocorrelation of 0.134 (t-stat of 4.48) for the Tradeoff measure, making its current level a good instrument for its future realization, which according to our model, is a forecast of LV anomaly return.

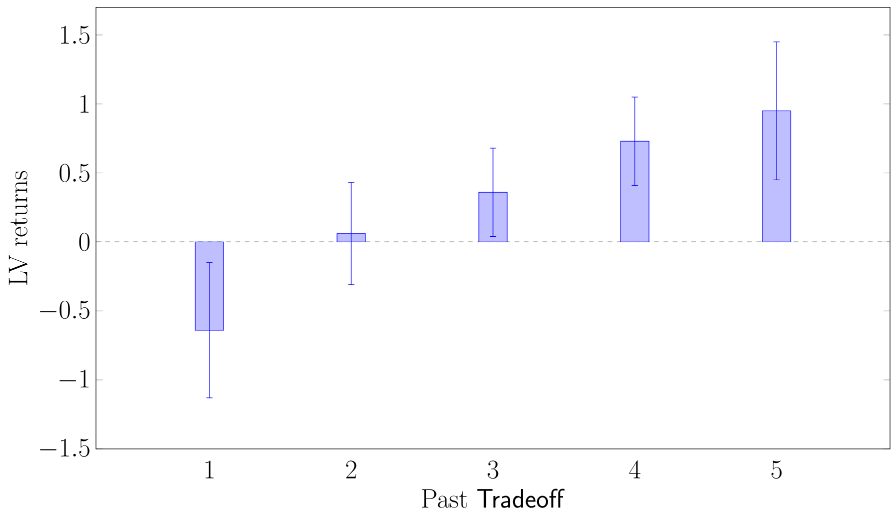
Our second candidate to proxy aggregate variance return trade-off of common factors is the isolated time-series bet of section 5.1. Statistics of this component which we denote by TS are reported in Table 5. If volatility-neutralizing is effective in removing all time-series effects from the LV strategy, then the TS component is a model-free proxy for factor trade-offs. TS can be interpreted as a catch-all proxy for the variance return trade-off of all observed and omitted common risk factors. This is our second predictor of LV returns. Although TS is created using LV returns and Tradeoff is created using factor data, their means and standard deviations are strikingly close with 0.11 and 4.40 for mean and standard deviation of Tradeoff, respectively, and 0.09 and 4.72 for those of TS. The measures have a large pairwise correlation coefficient of 0.58.

Figure 3 shows a predictive time-series sort of LV returns based on each measure. We use a total sample time-series sort to create five subsamples. Each month is categorized based on the level of Tradeoff (panel a) or TS (panel b) into five groups. We then compute the average next month return of the LV strategy for each group. Figure 3 shows that the LV strategy earns monotonically higher expected returns as factor trade-offs weaken. LV loses −0.64% following months of strong factor trade-offs (quintile 1) and earns 0.95% following weakest trade-off regimes. This large spread in returns can be contrasted with the strategy's unconditional return of only 0.28%. The alternative TS measure produces an even larger spread. LV earns −0.80% in months followed by lowest values of TS and an average of 1.10% in months followed by the highest.

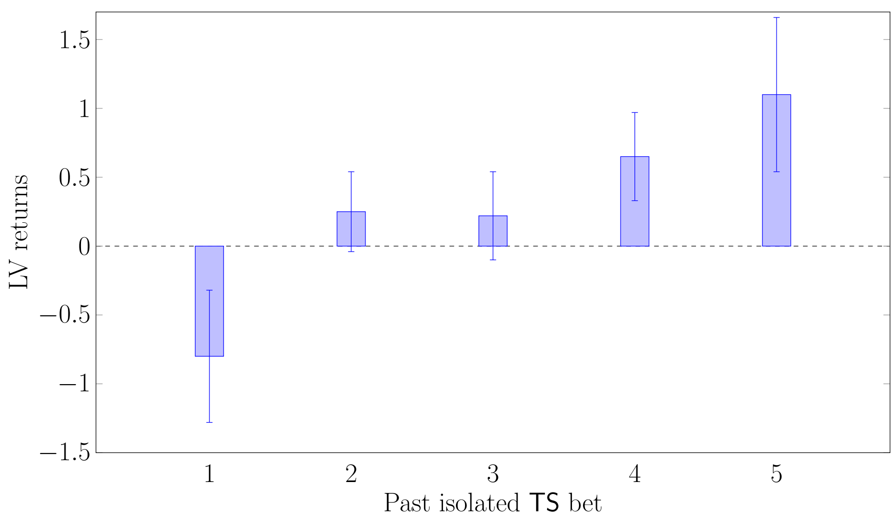
The monotonic pattern in Figure 3 is entirely consistent with the linear factor model representation of variance profits that relates variation in LV returns to variation in intertemporal common factor trade-offs. Our empirical measure of the intertemporal variance return trade-off is persistent and transmits into the cross-section. Under these joint conditions, the strength of trade-off this month predicts next month pricing of variance in the cross-section.

We run predictive regressions on LV returns. Campbell and Thompson (2007) show that the relative size of R^2 of a predictive regression with respect to the squared of Sharpe ratio of the strategy reflects the value of information in a signal for improving expected returns. A univariate predictive regression thus gives us an upper bound on the potential improvement from timing LV using the regressor. Models (1) and (4) of Table 6 report the results from running two univariate regressions of expected LV returns on Tradeoff and TS. Both measures are significantly related to expected returns with large t-statistics. These predictive models produce R^2 of 1.89% and 1.33% that are large relative to LV's monthly squared Sharpe ratio of 0.37%, suggesting that potential benefits of timing LV based on common factor trade-offs is noteworthy.

It is critical to verify the specifics though which factor trade-offs predict LV returns. It can be that aggregate trade-off predicts LV indirectly through market returns. That is, if a low risk



(a) Predicting LV using factor trade-offs



(b) Predicting LV using its isolated time-series component

Figure 3: Next month return of the low variance strategy across Tradeoff **subsamples.**
The top panel of this figure presents average next-month returns for the LV strategy across five subsamples sorted by Tradeoff. Each month is sorted into one of the five subsamples based on its level of Tradeoff. Each subsample includes 132 months. Within each subsample, we compute the time-series average of next month returns for each cross-sectional strategy. The bottom panel presents average next-month returns for the LV strategy across five subsamples sorted by TS. The rest of the procedure is similar to Panel a.

return trade-off predicts lower market returns, and market returns are negatively associated with LV returns, then our measures association with LV returns may be a manifestation of its relationship with market. We investigate whether Tradeoff tell us anything beyond what can be learned from future market returns by controlling for contemporaneous market returns. This is a strict robustness test of Tradeoff redundancy or lack thereof because market is the natural driver of risk-based strategy returns. Results in models (2) and (5) show that both Tradeoff and TS remain highly significant with *t*-stats of 3.04 and 2.82 after controlling for market returns.

In columns (3) and (6), we test whether factor trade-offs predict LV after controlling for contemporaneous returns of *all* five factors. This test of alpha predictability is an extremely strict robustness check because it controls for contemporaneous returns of factors whose lagged data are the inputs to our predictor. The coefficients in columns 3 and 6 are smaller but still significant with *t*-statistics of 2.20 and 1.96, indicating that common factor trade-offs contain information for predicting LV returns that is not present in future returns of the factors themselves. These results could only be obtained if two dynamics jointly and strongly hold in data: a) the variance return trade-off in time-series takes the form of variance return trade-off of the cross-section and b) factor trade-offs are persistent. We find compelling evidence in support of this collective existence.

5.3.2 An out-of-sample strategy

A natural question is whether Tradeoff performs well as an out-of-sample predictor and thus a tool to time the low variance strategy. The in-sample predictability may not transform into profitability of a trading strategy if the unconditional return difference between low variance and high variance stocks is large enough to overwhelm our ability to forecast returns. To explain it in the context of Campbell and Thompson (2007), the size of out-of-sample R^2, rather than in sample R^2, determines information value.

One way to implement a trading strategy is to use an initial subperiod of the sample to "train" our measures, after which we fit predicted values to LV returns recursively by adding one observation at a time until the sample is exhausted. An alternative approach that is easier and model free is to use the sign of our signal (factor trade-offs) to trade on or against variance in the cross-section. We lose some information by using a binary predictor instead of its continuous form but sign is an appealing predictor given its simplicity. It turns out that our simple measure, sign of Tradeoff (or TS), contains considerable amount of information about the future price of risk in the cross section. The standard LV strategy that buys low variance stocks and shorts high variance stocks earns a monthly premium of 0.28% (t-stat $= 1.54$). More than all of this premium is realized in months following negative trade-off regimes (0.97%, t-stat $= 2.65$). In other months, when the risk return trade-off is expected to be positive, risky stocks outperform and the strategy earns -0.21% (t-stat $= -0.81$). From a portfolio choice perspective, this result implies that a mean-variance investor may time exposure to cross-sectional variance according to the intertemporal trade-offs.

We examine whether a real-time investor would have benefited from timing the LV strategy by following a risk-on or risk-off (RoRo) strategy according to the past sign of the intertemporal risk return trade-off using this simple trading program:

1. Compute Tradeoff or TS every month.

2. Risk-on (buy the high variance and sell the low variance) at $t + 1$ if the time t trade-off is larger than the unconditional mean of 0.28%, and risk-off (buy the low variance and sell the high variance) otherwise.[17]

This implies that the scaling parameter is -1. The investor invests in either LV or its negative based on the real-time signal. If variance return trade-offs of common factors in time-series correctly identify whether variance commands a premium in the cross section ex-ante, the RoRo strategy will be superior to the static version.

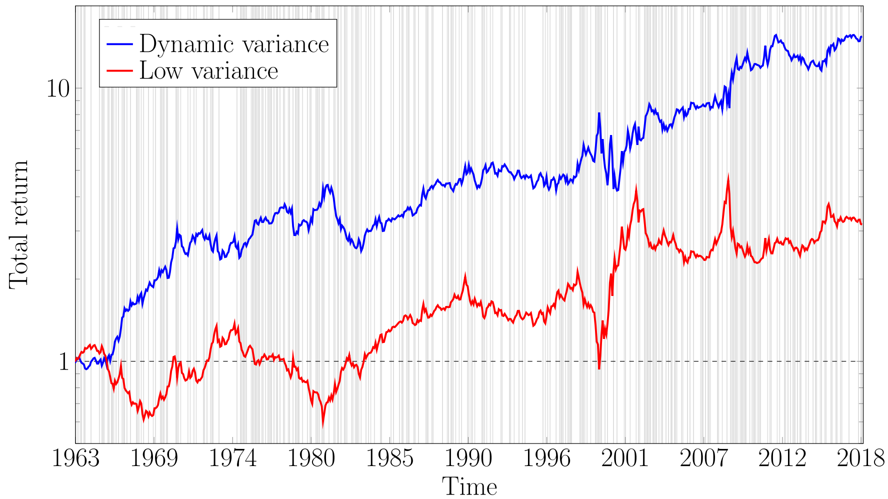
Figure 4 compares the performance of two Risk-on Risk-off (RoRo) strategies to that of the static LV. A unit invested in the static low-variance strategy becomes 3.2 by the end of the sample period. Both dynamic strategies see steady gains during the whole sample period and earn

[17] Improvements presented here are robust to a large range of assumptions for the unconditional mean, from zero to $+0.50\%$.

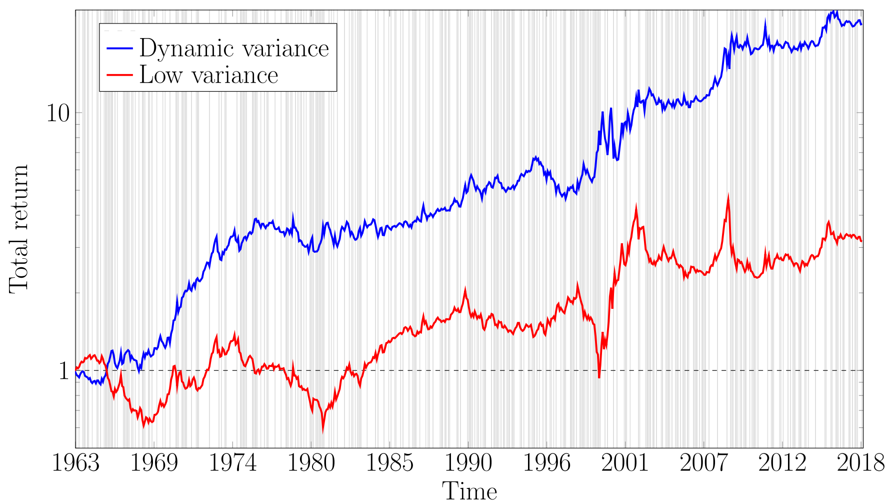
total returns of 15.5 for the strategy that uses Tradeoff as its signal (Panel a) and 22.1 for the strategy that trades using TS (Panel b). The substantial discrepancy between the performance of our RoRo strategies and the static LV comes from months, highlighted in gray, in which our dynamic strategies deviate from the unconditional strategy and buy (rather than sell) high risk stocks. Regarding average returns, the dynamic strategies based on Tradeoff and TS earn a monthly premium of 0.53% and 0.58% compared to 0.28% of the standard LV. All three strategies have similar standard deviations of about 4.70%.

The static low variance strategy delivers reasonable returns up to 09/2009 after which its total returns are close to zero. At the first glance, this suggests that market has been pricing volatility with less mispricing through time, a result that can be interpreted as an improvement in the efficiency of the overall market. It can also be a reflection of the post publication decline in anomaly premiums or data snooping biases (McLean and Pontiff, 2016; Linnainmaa and Roberts, 2018). In our context, the root cause of this degradation in performance is more frequent realizations of positive intertemporal risk return relationship in recent data, which translates into positive risk return trade-off in the cross-section. Consistent with this interpretation, I find that the earlier period (prior to 09/2009) when the LV strategy earns a high average of 0.36% is accompanied by larger negative intertemporal factor trade-offs. The average Tradeoff is -0.13 per month during this time period. After 09/2009, we have 201 months in the sample wherein the average variance return trade-offs is larger at -0.01 and LV earns only 0.12% per month. We conclude that the underperformance of the LV strategy in recent data is at least partly because the intertemporal variance return relationship is not as negative as the earlier period. Further supporting this interpretation, our timed strategies—that should be immune to the status of intertemporal trade-offs—earn about 0.55% per month in both subsamples.

Whether the dynamic strategy continues to enhance the returns of the static LV in future depends on the stability of the features underlying its profitability in the first place. The success of the strategy will for the most part rest on time-varying and autocorrelated variance return



(a) Timing LV using factor trade-offs



(b) Timing LV using its isolated time-series bet

Figure 4: Dynamic vs static low-variance strategy
This figure presents compounded returns for the dynamic risk-on risk-off (RoRo) and unconditional low-variance strategies. The risk-off strategy always sells high variance and buys low variance. RoRo buys high variance and sells low variance if risk factor trade-off is positive, and buys low variance and sells high variance otherwise. Factor trade-offs are measured in two different ways. Periods during which the RoRo of each strategy buys high variance and sells low variance are grey. Both strategies are fully out-of-sample.

trade-offs of common factors. Without the time-varying feature, the strategy will not be any different from a passive strategy. It is hard to see an economic mechanism that might cause such a break considering the stochastic nature of market returns and variance. The key to the success of the strategy is thus the continual of the persistence property in factor returns and variances. Weakened or a structural break in autocorrelations will restrict the capability of Tradeoff to predict the future price of risk in the cross-section.

We test if exposure to standard risk factors can explain the returns to our RoRo strategies. In Table 7, we first regress the RoRo strategy developed using Tradeoff on its static version, i.e., the standard LV. This model does not do much in explaining the premium indicated by the insignificant coefficient on LV and a zero R^2, meaning that the dynamic strategy does not correlate with its static version. We next regress the strategy on the BAB factor. The strategy loads negatively on the BAB factor and the intercept (α) increases to 0.65%. Controlling for market in the third model has a similar effect on the intercept. In the fourth model, we control for the Fama-French five factor model. The RoRo strategy does not load much on any factor and as a result the model produces a small R^2 of only 1.62%. Augmenting the Fama-French five factor model with a momentum factor in model (5) does not make a large difference. The same inference carries on to regressions (6) through (10) with the alternative RoRo strategy as the dependent variable. We conclude that we are capturing a totally different effect.

6 An indirect test of the multi-factor model

This section considers how well variance return trade-offs of alternative subsets of five factors relate to the pricing of variance in cross-section. Because a collection of all risk factor variance return trade-offs predicts LV returns, the trade-off measures that consist of a larger number of "true" factors should have more power in describing LV returns. Putting it on its head, if we abide by our framework, this test can be interpreted as an indirect test of the true factors that underline the return process. In this case, the predictive power of various trade-off measures provides evidence on the better asset pricing model.

Our methodology is a three-stage process: 1) we construct all possible measures of aggregate Tradeoff for each set from one to the full set of five factors 2) each Tradeoff is used in a univariate predictive regression of the form $LV_{t+1} = \alpha + \beta \, \mathsf{Tradeoff}_t + \epsilon$ to forecast LV returns: 3) we record these estimates and average their values and t-statistics for each set.[18] The addition of every risk factors should improve our forecast power if the factors represent the return generating process well and exhibit the autocorrelation property.

Table 8 shows that the coefficient of Tradeoff increases from 0.15 when the measure is constructed using only one factor to 0.22 when all factors are included. Average t-stats increase by more than 70% from 2.08 to above 3 for more refined measures that alternate between more factors. Moreover, the improvements do not flatten out and it appears that adding more factors (beyond our five) to the set may further increase forecast accuracy. The impact of additional factors in forecast accuracy is more evident when the dependent variable is the risk-adjusted LV returns. The results in the last column show that the significant Tradeoff for predicting LV's alpha are those that contain four or more factors. Altogether, we find that using multiple factors for estimating the aggregate trade-off measure produce remarkable improvement in the economic performance of the predictor consistent with the notion that a collection of variance return trade-offs of risk factor in time-series command the price of variance in the cross-section.

7 How do trade-offs spread into the cross-section?

We test the mechanism through which risk return trade-off in time-series carries over into the cross-section. If a common component (factor trade-offs) drives all of our findings, we should find the effect in every corner of the cross-section but more from volatile stocks since they are the main vehicles that transmit trade-offs into the cross-section. We examine whether a common component drives our results by estimating the following regression for all decile portfolios sorted on variances:

$$r_t = a + b\,\mathsf{Tradeoff}_{t-1} + \epsilon_t$$

[18] For example, for the set of two factors, we first select two factors (i.e., MKTRF and SMB), compute equation 14 and predict the next month return of LV in a univariate regression. We record the coefficient and t-stat of the regression. This is one repetition. We then replace SMB with another factor, compute the measure, and rerun the regression. With replacement in each iteration we examine the robustness of the measure to a different composition. We conduct this exercise for $\binom{5}{2} = 10$ repetitions to find estimates of all possible sets consisting of two factors. For three, four, and five factors, we repeat this exercise 10, 5, and 1 times, respectively.

where r is the return on a portfolio and $\mathsf{Tradeoff}_{t-1}$ is as before, the *negative* of risk return trade-off constructed using all 5 factors. The blue bars in Figure 5 plot the estimates and confidence intervals for estimates of $\hat{b}$. Consistent with the model's prediction, the impact of factor trade-offs on expected returns is monotonic in portfolio variance. High variance assets are the main vehicles that magnify and scatter factor trade-offs into the cross-section. This pattern is far from random and confirms the robustness of Tradeoff in forecasting the returns to variance portfolios. Tradeoff predicts return spreads between any two portfolios whose variances are systematically different.

We run two more regressions for each portfolio. One controls for contemporaneous market returns which can be interpreted as a predictive model of CAPM alpha (red bars), and another model controls for all five factors, which is a prediction of FF5 alpha (presented in green bars). These estimates show that the general conclusions drawn from regression on raw returns hold when Tradeoff predicts CAPM or FF5 risk-adjusted returns. A negative trade-off in time series is associated with risk-adjusted *outperformance* of low risk stocks, *average performance* for mid-variance portfolios, and large *underperformance* of volatile portfolios.

8 Conclusion

In this paper, we show that the evolution of price of variance in cross-section is linked to the pricing of the variance in time-series. Cross-sectional low-risk strategies are by definition time-varying bets against the variance of common risk factors. Any change in factor portfolio variances during the portfolio formation period will be reflected in strategy's factor exposure during the investment period, implying that returns to the low variance strategy contains all time-series effects arising from factor's stochastic volatility. I show that arbitrage activity declines during times when a bet on LV lines-up with factor variances, consistent with the hypothesis that risk-averse rational traders are reluctant to arbitrage the anomaly away because the potential profits come at substantial factor risks. Further supporting this hypothesis, factor risk is

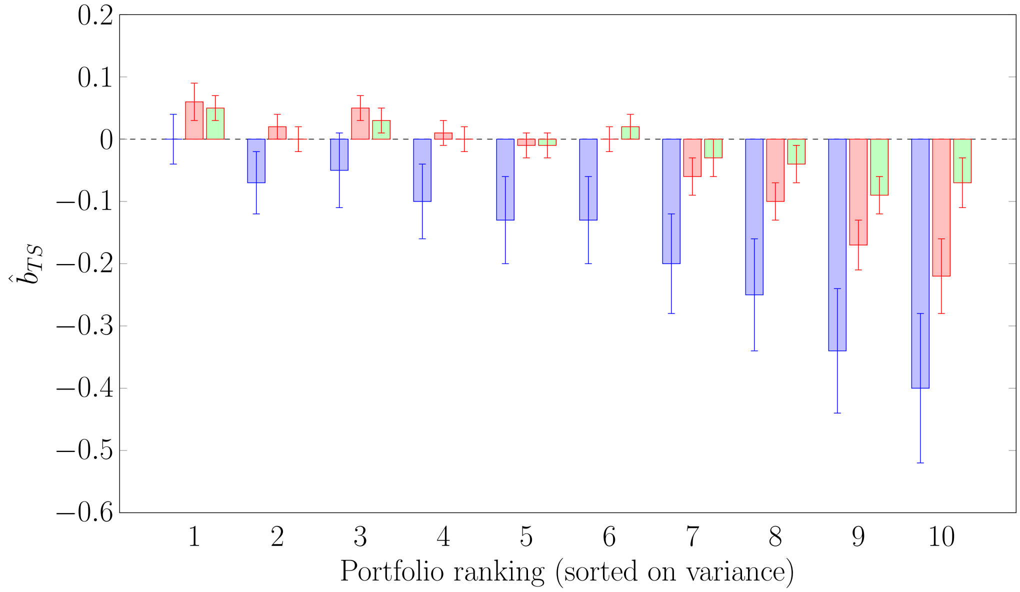


Figure 5: How Tradeoff **spreads into the cross-section.**
The blue bars in this figure presents point estimates and their confidence intervals obtained from regressions of returns of portfolio sorted on variance on lagged Tradeoff. The red bars present the slopes for the model that controls for contemporaneous market returns. The green bars present slopes for another model that controls for contemporaneous returns of all five factors.

associated with higher future alpha to the anomaly. In fact, the low variance strategy is not an anomaly during low factor risk periods.

I find that the component of the strategy that stems from dynamic factor exposure produces much of its variability. However, this component can be predicted using its underlying source, i.e., the variance return trade-off of common risk factors in time-series. A strong or weak trade-off in risk factors this month is more likely to be followed by a trade-off of the same type next month, which then spreads to the cross-section and impacts low variance profits. I develop strategies that take advantage of this persistence in time-series trade-off to predict the return on the variance anomaly. Our strategy bets *against* variance in the cross-section if the expected time-series trade-off is negative, and bets *on* variance otherwise. The dynamic strategy outperforms its unconditional low-risk equivalent significantly and produces large alphas with respect to several asset pricing models. The strategy owes its success to the persistence in intertemporal risk return trade-off derived from autocorrelations in factor returns and variance.

The future performance of the strategy will depend on the perseverance of the autocorrelation property of the first and second moment of common factor returns.

References

Abel, A. B. (1988). Stock prices under time-varying dividend risk: An exact solution in an infinite-horizon general equilibrium model. *Journal of Monetary Economics*, 22(3):375–393.

Ang, A., Hodrick, R. J., Xing, Y., and Zhang, X. (2006). The cross-section of volatility and expected returns. *The Journal of Finance*, 61(1):259–299.

Ang, A., Hodrick, R. J., Xing, Y., and Zhang, X. (2009). High idiosyncratic volatility and low returns: International and further U.S. evidence. *Journal of Financial Economics*, 91(1):1–23.

Antoniou, C., Doukas, J. A., and Subrahmanyam, A. (2015). Investor sentiment, beta, and the cost of equity capital. *Management Science*, 62(2):347–367.

Arnott, R. D., Clements, M., Kalesnik, V., and Linnainmaa, J. T. (2018). Factor momentum. *Available at SSRN: https://ssrn.com/abstract=3116974*.

Backus, D. K. and Gregory, A. W. (1993). Theoretical relations between risk premiums and conditional variances. *Journal of Business & Economic Statistics*, 11(2):177–185.

Baker, M., Bradley, B., and Wurgler, J. (2011). Benchmarks as limits to arbitrage: Understanding the low-volatility anomaly. *Financial Analysts Journal*, 67(1):40–54.

Bali, T. G., Brown, S. J., Murray, S., and Tang, Y. (2017). A lottery-demand-based explanation of the beta anomaly. *Journal of Financial and Quantitative Analysis*, 52(6):2369–2397.

Bali, T. G., Cakici, N., and Whitelaw, R. F. (2011). Maxing out: Stocks as lotteries and the cross-section of expected returns. *Journal of Financial Economics*, 99(2):427–446.

Barberis, N. and Huang, M. (2008). Stocks as lotteries: The implications of probability weighting for security prices. *American Economic Review*, 98(5):2066–2100.

Barroso, P. and Maio, P. (2018). Managing the risk of the beta anomaly. *Available at SSRN 2876450*.

Black, F. (1972). Capital market equilibrium with restricted borrowing. *The Journal of Business*, 45(3):444–

455.

Brandt, M. W. and Kang, Q. (2004). On the relationship between the conditional mean and volatility of stock returns: A latent var approach. *Journal of Financial Economics*, 72(2):217–257.

Campbell, J. Y. and Thompson, S. B. (2007). Predicting excess stock returns out of sample: Can anything beat the historical average? *The Review of Financial Studies*, 21(4):1509–1531.

Cederburg, S., O'Doherty, M. S., Wang, F., and Yan, X. S. (2019). On the performance of volatility-managed portfolios. *Journal of Financial Economics (JFE), Forthcoming*.

Christoffersen, S. E. and Simutin, M. (2017). On the demand for high-beta stocks: Evidence from mutual funds. *The Review of Financial Studies*, 30(8):2596–2620.

Conrad, J. and Kaul, G. (1998). An anatomy of trading strategies. *Review of Financial studies*, 11(3):489–519.

Ehsani, S. and Linnainmaa, J. (2021). Factor momentum and the momentum factor. *Forthcoming, Journal of Finance*.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, pages 987–1007.

Fama, E. F. and French, K. R. (1992). The cross-section of expected stock returns. *The Journal of Finance*, 47(2):427–465.

Fama, E. F. and French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, 116(1):1 – 22.

Fleming, J., Kirby, C., and Ostdiek, B. (2001). The economic value of volatility timing. *The Journal of Finance*, 56(1):329–352.

Frazzini, A. and Pedersen, L. H. (2014). Betting against beta. *Journal of Financial Economics*, 111(1):1–25.

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.

Han, B. and Kumar, A. (2013). Speculative retail trading and asset prices. *Journal of Financial and Quantitative Analysis*, 48(2):377–404.

Harvey, C. (1989). Time-varying conditional covariances in tests of asset pricing models. *Journal of Financial Economics*, 24(2):289–317.

Herskovic, B., Kelly, B., Lustig, H., and Van Nieuwerburgh, S. (2016). The common factor in idiosyncratic volatility: Quantitative asset pricing implications. *Journal of Financial Economics*, 119(2):249–283.

Hong, H. and Sraer, D. A. (2016). Speculative betas. *The Journal of Finance*, 71(5):2095–2144.

Kozak, S., Nagel, S., and Santosh, S. (2018). Interpreting factor models. *The Journal of Finance*, 73(3):1183–1223.

Lehmann, B. N. (1990). Fads, martingales, and market efficiency. *The Quarterly Journal of Economics*, 105(1):1–28.

Lettau, M. and Ludvigson, S. C. (2009). *Handbook of Financial Econometrics: Tools and Techniques*, volume 1, chapter 11 - Measuring and Modeling Variation in the Risk-Return Trade-off, pages 617–690. Elsevier.

Linnainmaa, J. T. and Roberts, M. R. (2018). The history of the cross-section of stock returns. *The Review of Financial Studies*, 31(7):2606–2649.

Lo, A. W. and MacKinlay, A. C. (1990). When are contrarian profits due to stock market overreaction? *Review of Financial Studies*, 3(2):175–205.

Lou, D. and Polk, C. (2012). Comomentum: Inferring arbitrage activity from return correlations. In *Available at SSRN: https://ssrn.com/abstract=2023989*.

Lundblad, C. (2007). The risk return tradeoff in the long run: 1836–2003. *Journal of Financial Economics*, 85(1):123–150.

McLean, R. D. and Pontiff, J. (2016). Does academic research destroy stock return predictability? *The Journal of Finance*, 71(1):5–32.

Merton, R. C. (1973). An intertemporal capital asset pricing model. *Econometrica: Journal of the Econometric Society*, 41(5):867–887.

Merton, R. C. (1980). On estimating the expected return on the market: An exploratory investigation.

Journal of Financial Economics, 8(4):323–361.

Moreira, A. and Muir, T. (2017). Volatility-managed portfolios. *The Journal of Finance*, 72(4):1611–1644.

Novy-Marx, R. and Velikov, M. (2019). Betting against betting against beta. *Available at SSRN 3300965*.

Pontiff, J. (2006). Costly arbitrage and the myth of idiosyncratic risk. *Journal of Accounting and Economics*, 42(1-2):35–52.

Schneider, P., Wagner, C., and Zechner, J. (2017). Low risk anomalies? *Available at SSRN: https://ssrn.com/abstract=2593519 or http://dx.doi.org/10.2139/ssrn.2593519*.

Stambaugh, R. F., Yu, J., and Yuan, Y. (2015). Arbitrage asymmetry and the idiosyncratic volatility puzzle. *The Journal of Finance*, 70(5):1903–1948.

Whitelaw, R. F. (1994). Time variations and covariations in the expectation and volatility of stock market returns. *The Journal of Finance*, 49(2):515–541.

A Appendix

A.1 Decomposing variance strategy profits

We derive the profits of the linear weighting variance strategy. The amount invested in asset p at time t is determined by its $t-1$ variance relative to the average:

$$w_{p,t} = \mathsf{var}_{p,t-1} - \overline{\mathsf{var}_{t-1}}$$

The return of a strategy that trades asset p is:

$$V^p = \left(\mathsf{var}_{p,t-1} - \overline{\mathsf{var}_{t-1}}\right) r_{p,t}$$

where $r_{p,t}$ is the realized return of asset p at time t. The return for the strategy that trades all assets in the cross-section becomes:

$$\begin{aligned}
&\mathbb{E}[V^{CS}] = \mathbb{E}\big[\frac{1}{P}\sum_{s=1}^{P}\big(\mathsf{var}_{p,t-1} - \overline{\mathsf{var}_{t-1}}\big)\, r_{p,t}\big] \\
&= \frac{1}{P}\sum_{p=1}^{P}\mathbb{E}\big[\mathsf{var}_{p,t-1}\, r_{p,t}\big] - \frac{1}{P}\sum_{p=1}^{P}\mathbb{E}\big[\overline{\mathsf{var}_{t-1}}\, r_{p,t}\big] \\
&= \frac{1}{P}\sum_{p=1}^{P}\mathbb{E}\big[\mathsf{var}_{p,t-1}\big]\,\mathbb{E}\big[r_{p,t}\big] + \frac{1}{P}\sum_{p=1}^{P}\mathsf{Cov}\big(\mathsf{var}_{p,t-1}, r_{p,t}\big) - \frac{1}{P}\sum_{p=1}^{P}\mathbb{E}\big[\overline{\mathsf{var}_{t-1}}\big]\,\mathbb{E}\big[r_{p,t}\big] - \frac{1}{P}\sum_{p=1}^{P}\mathsf{Cov}\big(\overline{\mathsf{var}_{t-1}}, r_{p,t}\big) \\
&= \frac{1}{P}\sum_{s=1}^{P}\sigma_p^2\,\mu_p + \frac{1}{P}\sum_{s=1}^{P}\mathsf{Cov}\big(\mathsf{var}_{p,t-1}, r_{p,t}\big) - \frac{1}{P}\sum_{s=1}^{P}\overline{\sigma^2}\,\mu_p - \mathsf{Cov}\big(\overline{\mathsf{var}_{t-1}}, \overline{r_t}\big) \\
&= \frac{1}{P}\sum_{p=1}^{P}\mathsf{Cov}\big(\mathsf{var}_{p,t-1}, r_{p,t}\big) - \mathsf{Cov}\big(\overline{\mathsf{var}_{t-1}}, \overline{r_t}\big) + \frac{1}{P}\sum_{p=1}^{S}\big(\sigma_p^2 - \overline{\sigma^2}\big)\big(\mu_p - \overline{\mu}\big),
\end{aligned}$$

where the last line is derived by adding $\frac{1}{P}\sum_{p=1}^{P}(-\sigma_p^2\,\overline{\mu} + \overline{\sigma^2}\,\overline{\mu})$, a zero term.

A.2 Data for estimating the profit matrix

Our candidate assets for implementing the trading strategy are decile equity portfolios sorted on variance. We need both pre-and post-ranking data for each portfolio to derive the profit matrix. The sample period is from 1963/07 to 2018/06 and consists of all stocks in the CRSP database with shares codes of 10 and 11. For each stock with at least twelve daily observations, we compute variance using a month worth of daily data. This information is known at the end of month t and we use this as conditioning information in forming portfolios for the next month. Stocks are then sorted into 10 decile portfolios based on an ascending sort of variance. Pre-ranking daily portfolio returns for month t, and post-ranking portfolio returns for the following month (month $t+1$) are computed, after which the estimation and formation procedures are repeated until the sample's last month. Pre-ranking daily returns and post-ranking monthly returns for each portfolio are then linked. The time-series of daily pre-ranking and monthly post-ranking returns of 10 decile portfolios are the test assets used to implement the strategy and estimate its sources of profits.

Table A1 presents descriptive statistics for the variance sorted portfolios.[19] The monotonic pattern in pre-ranking standard deviations continues to the next month indicating that volatility is persistent and past volatility of a portfolio is a good proxy for its future volatility. Pre- and post-ranking portfolio betas demonstrate very similar pattern which highlights the strong relation between risk measures. The last column displays the well-known puzzle. Average returns are about the same for low-variance portfolios but drop steeply from portfolios 5 to 10 such that the most volatile portfolio earns negative returns. The CAPM alphas, reported in the last column, are even more anomalous with the most volatile portfolio earning a risk-adjusted return of -1.55% per month with a t-stat of -6.09.

[19] We present standard deviation rather than variance for ease of interpretations. Our tests use variance.

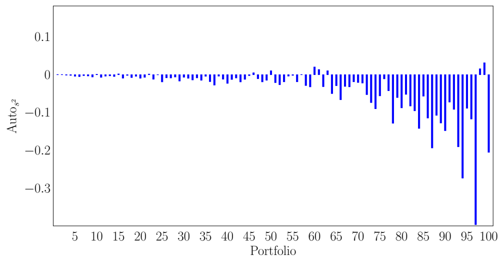
A.3 Time-series effects of systematic and residual variance for prcentile portfolios

Aggregating many stocks into portfolios eliminates most of the variation in the firm-specific component of returns because each decile portfolio contains on average around 400 stocks and is well-diversified. The trivial role of residual risk in creating the time-series effects of Table 4 may thus pertain to the choice of our test assets. This concern can be alleviated by forming a larger cross-section of variance-sorted portfolios. Instead of decile portfolios, we sort stocks into percentile portfolios and compute the pre- and post-ranking returns for each portfolio. We then estimate the covariance matrix of equation 5.[20]

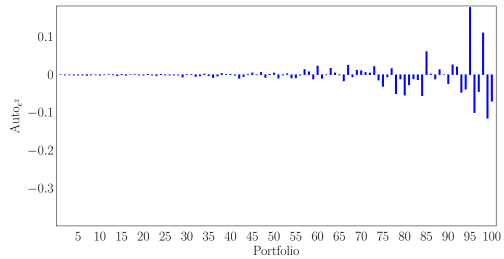
Figure A1 presents the profit break down by each component at the percentile portfolio level. The systematic and idiosyncratic effects for each component are presented in the same row and on the same scale for ease of comparison. Graphs A1a and A1b present the size of own-covariance between systematic variance and expected returns, and residual variance and expected returns, respectively. The negative bars in Graphs A1a imply that high systematic variance is followed by lower returns for most portfolios. Similar to previous tests on decile portfolios, the negative own-covariance effect persists in the cross-section and is largest among HV portfolios. Graph A1b shows that residual variance does not covary with expected returns, consistent with the results in Table 4.

We plot $-1\times$Cross-covariance in Graphs A1c, the positive bars are thus representative of negative cross-covariances, indicating that high systematic volatility of one portfolio signals lower returns on other portfolios, mitigating the negative volatility effect. we plot the cross-covariances associated with residual variance in Graph A1d. Residual variance of a portfolio does not predict low or high returns on other assets, the impacts switch frequently and randomly across portfolios, once again confirming that residual risk does not play a role in producing time series relationships between volatility and returns.

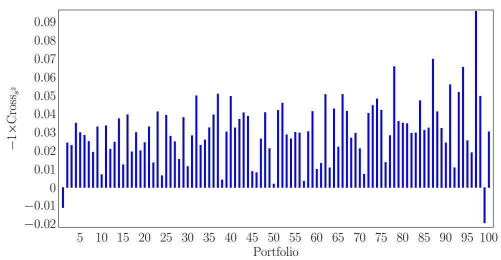
[20] Implementing the linear weighting strategy on the percentile portfolios produces a *t*-stat of -1.96 compared to the -1.19 for the strategy on decile portfolios.



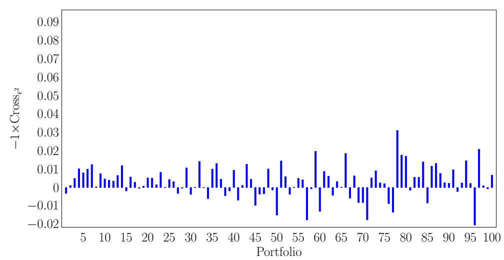
(a) Own (systematic)



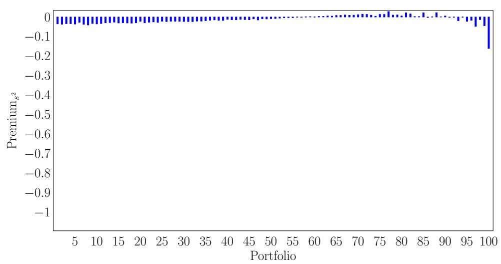
(b) Own (residual)



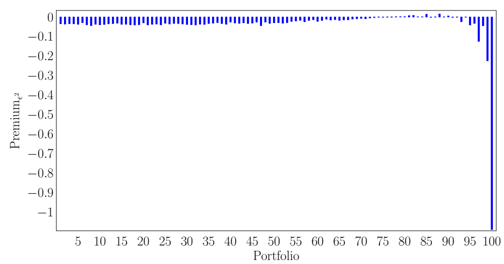
(c) Cross (systematic)



(d) Cross (residual)



(e) Premium (systematic)



(f) Premium (idiosyncratic)

Figure A1: Profit breakdown of each component by each portfolio.
We sort stocks into one hundred portfolios based on variance. We then form the variance strategy and compute its components using the decomposition matrix of equation 5. This graph shows the contributions of each percentile portfolio to each of the six component.

The bottom two graphs (Graphs A1e and A1f) display portfolio contributions deriving from unconditional dispersion in returns. The negative bars indicate that unconditional bets on both systematic and idiosyncratic variances earn negative premiums in most of the cross-section. Also, the negative bars are larger in graph A1f compared to graph A1e, showing that residual variance plays a more prominent in producing the unconditional component of the low-risk anomaly. In summary, the results using percentile portfolios are entirely consistent with those of Table 4. Systematic variance is producing all the time-series patterns between past variance and expected returns, whereas both systematic and residual variance command unconditional anomalous differences in returns.

A.4 The intertemporal risk return trade-off

Merton (1973) derives a linear relation between market excess returns, its conditional variance, and its conditional covariance with the state variable F

$$\mathbb{E}[r_{M,t+1}] = \frac{-J_{WW}W}{J_W}\mathsf{var}(r_{M,t}) + \frac{J_{WF}}{J_W}\mathsf{Cov}\big(r_{M,t}, F_t\big) \tag{15}$$

where $J(W(t), F(t), t)$ is the utility of wealth function, $W(t)$ is wealth, and $F(t)$ is the state variable F that describes the changes in the investment opportunity set. The term $\frac{-J_{WW}W}{J_W}$ is linked to the price of risk or the coefficient of relative risk aversion which we denote by γ. Assuming the investment opportunity set is static $\big(\mathsf{Cov}(r_{M,t}, F_t) = 0\big)$, the model reduces to this univariate form:

$$\mathbb{E}[r_{M,t+1}] = \gamma\,\mathsf{var}(r_{M,t}) \tag{16}$$

Literature typically assumes γ to be positive, but a lack or even negative γ is not necessarily against theory. Abel (1988) and Backus and Gregory (1993) provide alternative theories in which market risk premium is lower during riskier periods. If investors save more when risk increases, and all assets available for investment carry risk (no real risk-free investment is available), then prices of risky assets may bid up significantly during riskier periods, reducing the risk premium. In this framework, the shape of the relation between the conditional variance and expected returns is unrestricted with increasing, decreasing, or flat patterns all possible. Thus, unlike the cross-section in which the risk return trade-off is expected to be positive, whether such relation should exist between risk and return in time-series remains an open question.

Numerous empirical studies explore the intertemporal relationship between risk and return. This literature is too extensive to review comprehensively, we briefly discuss those that relate most to our framework. Merton (1980), French et al. (1987), Harvey (1989), Glosten et al.

(1993) and much of the earlier literature employ variations of the restricted model to find γ, and find positive, negative, or flat estimates. Typically, these papers analyze the intertemporal risk-return relation by studying the time-series relation between the conditional mean and the conditional variance of market returns, neither of which is observed. As a result, the tests are very sensitive to model specification because in practice, papers impose strong modeling assumptions before estimating the shape of the relationship. Other studies show that the risk-return relation is counter-cyclical and varies considerably through time (Whitelaw, 1994; Brandt and Kang, 2004; Lundblad, 2007; Lettau and Ludvigson, 2009), suggesting that a generalized form in which the price of risk is time-varying (γ_t) better describes the intertemporal relationship between conditional variance and expected returns.

Table 1: Factor variances command future performance of low-variance anomaly.

Panel a shows the results from regressions of future returns to the LV strategy on past high-minus-low factor variances–which are commanders of future factor exposure of the strategy. σ^2_{MKTcs}, the independent variables, $\sigma^2_{high} - \sigma^2_{low}$, are the variance of a factor's high portfolio minus the variance of its low portfolio. σ^2_{MKTcs} is the cross-sectional variance of the market portfolio. We also report the results for the average factor risk, $\overline{\sigma^2}$, which is the average of the other five high-minus-low variances. Panel b is similar to panel but controls for the FF5 returns. The regressions in Panel A predict returns and the regressions in Panel B predict alphas.

Panel A. $LV_{t+1} = a + b\,(\sigma^2_{f,t,high} - \sigma^2_{f,t,low}) + e$	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\overline{\sigma^2}$	−1.56						
	(−3.93)						
σ^2_{MKTcs}		−0.05					−0.30
		(−0.24)					(−1.29)
$\sigma^2_S - \sigma^2_B$			−0.01				−0.05
			(−0.04)				(−0.20)
$\sigma^2_H - \sigma^2_L$				−0.25			−0.25
				(−2.00)			(−1.78)
$\sigma^2_R - \sigma^2_W$					−0.74		−0.67
					(−3.24)		(−2.27)
$\sigma^2_C - \sigma^2_A$						−0.97	−0.55
						(−3.45)	(−1.52)
Intercept	0.28	0.32	0.28	0.26	0.11	0.04	0.19
	(1.55)	(1.30)	(1.53)	(1.42)	(0.58)	(0.18)	(0.77)
adj. R^2	2.1%	−0.1%	−0.2%	0.5%	1.4%	1.6%	2.3%

Panel B. $LV_{t+1} = a + b\,(\sigma^2_{f,t,high} - \sigma^2_{f,t,low}) + b_{FF5}FF5_{t+1} + e$	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\overline{\sigma^2}$	−0.81						
	(−4.06)						
σ^2_{MKTcs}		−0.13					−0.22
		(−1.16)					(−1.93)
$\sigma^2_S - \sigma^2_B$			0.07				0.12
			(0.55)				(0.94)
$\sigma^2_H - \sigma^2_L$				−0.21			−0.24
				(−3.31)			(−3.35)
$\sigma^2_R - \sigma^2_W$					−0.29		−0.28
					(−2.50)		(−1.88)
$\sigma^2_C - \sigma^2_A$						−0.41	−0.21
						(−2.87)	(−1.17)
Control	FF5	FF5	FF5	FF5	FF5	FF5	FF5
Intercept	0.28	0.37	0.27	0.26	0.21	0.18	0.30
	(2.95)	(2.90)	(2.75)	(2.71)	(2.10)	(1.72)	(2.40)
adj. R^2	75.4%	74.9%	74.8%	75.2%	75.1%	75.1%	75.6%

Table 2: Factor risk and arbitrage activity.

Table shows pairwise correlations between factor variance spreads and arbitrage activity in LV strategy. The first two columns report correlations between variance spreads and next month arbitrage activity. The following two columns report the contemporaneous correlation between arbitrage activity and variance spreads. t-statistics are reported in brackets.

	Arbitrage activity measure: $CoVol$			
	Predictive		Contemporaneous	
	Detrended	Not detrended	Detrended	Not detrended
$\overline{\sigma^2}$	0.23	0.20	0.22	0.19
	[6.20]	[6.01]	[5.93]	[5.44]
σ^2_{MKTcs}	0.05	-0.01	0.07	0.00
	[1.36]	[-0.41]	[1.82]	[-0.07]
$\sigma^2_S - \sigma^2_B$	0.15	0.22	0.19	0.26
	[4.47]	[6.58]	[5.79]	[7.89]
$\sigma^2_H - \sigma^2_L$	0.12	0.17	0.12	0.16
	[4.31]	[6.43]	[2.90]	[4.57]
$\sigma^2_R - \sigma^2_W$	0.08	-0.02	0.06	-0.04
	[2.16]	[-0.43]	[1.28]	[-0.89]
$\sigma^2_C - \sigma^2_A$	0.12	0.11	0.07	0.06
	[2.78]	[2.76]	[1.46]	[1.42]

Table 3: Profit decomposition of the variance strategy.

Panel A shows the premium of the variance strategy, and the premiums emerging from each of its three components. The aggregate contributions of each component is computed using equation 4. Panel B breaks down the contributions by portfolio. The *t*-statistics are obtained by block bootstrapping.

Panel A: Aggregate contributions of each component

		Time series		Cross section
	Strategy	Own	$-1\times$Cross	Premium
Premium	−2.93	−4.49	3.24	−1.68
t-stat	[−1.19]	[−0.98]	[1.04]	[−3.70]
Net contribution	100%	153%	−111%	57%

Panel B: Portfolio level contributions of each component

		Time series		Cross section
Portfolio	Strategy	Own	$-1\times$Cross	Premium
1	−0.15	−0.03	0.31	−0.42
2	−0.19	−0.05	0.26	−0.41
3	−0.15	−0.08	0.26	−0.33
4	−0.07	−0.17	0.39	−0.29
5	−0.08	−0.08	0.21	−0.21
6	0.06	−0.11	0.27	−0.09
7	0.04	−0.32	0.31	0.05
8	−0.15	−0.88	0.56	0.17
9	−0.34	−0.80	0.30	0.15
10	−1.90	−1.98	0.37	−0.29

Table 4: Profit decomposition by systematic and residual variances.

Panel A shows the premium of the variance strategy, and the premiums emerging from each of its six components. The aggregate contributions of each component is computed using equation 5. Panel B breaks down the premium and contributions by portfolio. The t-statistics are obtained by block bootstrapping.

Panel A: Aggregate contributions of each component

	Time series				Cross section	
	Own *Systematic*	Own *Residual*	$-1\times$Cross *Systematic*	$-1\times$Cross *Residual*	Premium *Systematic*	Premium *Residual*
Premium (%)	−4.61	0.11	3.18	0.06	−0.90	−0.78
t-stat	[−1.13]	[0.12]	[1.14]	[0.13]	[−3.15]	[−4.85]
Net contribution	157%	−4%	−108%	−2%	31%	26%

Panel B: Portfolio level contributions of each component

	Time series				Cross section	
Portfolio	Own *Systematic*	Own *Residual*	$-1\times$Cross *Systematic*	$-1\times$Cross *Residual*	Premium *Systematic*	Premium *Residual*
1	−0.03	−0.01	0.28	0.03	−0.32	−0.10
2	−0.04	−0.01	0.25	0.01	−0.29	−0.12
3	−0.08	0.00	0.25	0.00	−0.22	−0.11
4	−0.16	−0.01	0.37	0.02	−0.17	−0.12
5	−0.09	0.01	0.23	−0.02	−0.10	−0.11
6	−0.10	−0.01	0.27	0.00	−0.01	−0.08
7	−0.37	0.05	0.31	0.00	0.09	−0.04
8	−0.91	0.04	0.52	0.04	0.15	0.01
9	−0.82	0.02	0.31	−0.01	0.11	0.04
10	−2.01	0.03	0.38	−0.02	−0.14	−0.15

Table 5: Volatility neutralized low variance strategy

Table shows the return to the standard low-minus-high variance, its volatility neutral version, and the time-series bet. The volatility neutralized strategy is constructed in spirit of the betting-against-beta factor (Frazzini and Pedersen, 2014) and attempts to hedge out the time-series bet from the standard low-minus-high variance.

	Standard LV	Time series bet (TS)	Volatility neutral (VN)
Mean	0.28	0.09	0.19
SD	4.73	4.40	0.87
t-stat	[1.56]	[0.55]	[5.71]

Table 6: Predicting LV returns using factor trade-offs.

This table tests how two measures of common factor trade-offs predict the returns to the low-minus-high variance strategy. Tradeoff is calculated using factor data only, and TS is the component of LV returns that reflect the return to LV reflecting from the status of common factor risk return trade-off. Both variables are lagged and measure the negative of common factor trade-offs. The coefficients show the impact of a negative factor trade-off on future LV returns. The table also tests whether Tradeoff and TS predict CAPM and FF5 alphas. t-statistics are in parenthesis. N is 659.

	Dependent variable: LV returns					
	(1)	(2)	(3)	(4)	(5)	(6)
Tradeoff_{t-1}	0.22	0.14	0.07			
	[3.56]	[3.04]	[2.20]			
TS_{t-1}				0.12	0.09	0.04
				[2.97]	[2.82]	[1.96]
CONTROL	–	MKTRF	FF5	–	MKTRF	FF5
R^2	1.89%	46.51%	75.22%	1.33%	46.41%	75.19%

Table 7: Factor model regressions for the dynamic variance strategy.

We run 5 time-series factor model regressions on two RoRo strategies. The first specification is a univariate regression of the RoRo returns on its static low-variance version. The second specification controls for the betting against beta (BAB) factor. The third model is a CAPM regression. The fourth model controls for five factors of Fama and French. The last model is the 5-factor model augmented with a momentum factor. t-statistics are reported in parenthesis.

	RoRo (Tradeoff)					RoRo (TS)				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
α	0.53	0.65	0.59	0.52	0.54	0.54	0.68	0.71	0.62	0.63
	[2.87]	[3.43]	[3.23]	[2.74]	[2.79]	[2.98]	[3.60]	[3.93]	[3.33]	[3.34]
LV	0.01					0.13				
	[0.26]					[3.34]				
BAB		−0.15					−0.13			
		[−2.55]					[−2.15]			
MKTRF			−0.12	−0.09	−0.09			−0.24	−0.20	−0.21
			[−2.85]	[−1.83]	[−1.89]			[−5.85]	[−4.40]	[−4.41]
SMB				0.02	0.02				0.02	0.02
				[0.30]	[0.33]				[0.35]	[0.36]
HML				−0.11	−0.12				−0.04	−0.05
				[−1.21]	[−1.31]				[−0.44]	[−0.51]
CMA				0.34	0.35				0.28	0.29
				[2.50]	[2.54]				[2.13]	[2.16]
RMW				−0.04	−0.03				−0.03	−0.02
				[−0.40]	[−0.32]				[−0.30]	[−0.25]
UMD					−0.03					−0.02
					[−0.57]					[−0.35]
Adj R^2	0.00%	0.83%	1.07%	1.62%	1.51%	1.52%	0.55%	4.81%	5.25%	5.12%

Table 8: Predicting LV returns using Tradeoff **constructed with fewer factors**

This table repeats the results of Table 6 for different combinations of risk factors. I form Tradeoff using all possible subsets of the 5 factors and run a predictive regression on the LV returns. The remaining is similar to 6. The sample contains 659 monthly observations from 1963/08 to 2018/06.

	Dependent variable: LV returns					
	Univariate		+MKTRF		+FF5	
Number of factors	$\overline{\hat{b}}$	$\overline{t_{\hat{b}}}$	$\overline{\hat{b}}$	$\overline{t_{\hat{b}}}$	$\overline{\hat{b}}$	$\overline{t_{\hat{b}}}$
1	0.15	[2.08]	0.11	[1.93]	0.04	[1.23]
2	0.16	[2.41]	0.13	[2.33]	0.05	[1.53]
3	0.18	[2.83]	0.14	[2.65]	0.06	[1.80]
4	0.21	[3.23]	0.14	[2.88]	0.07	[2.03]
5	0.22	[3.56]	0.14	[3.04]	0.07	[2.20]

Table A1: Summary statistics for the test assets.

The table shows summary statistics for the test assets in this section. The test assets are decile portfolios sorted on total variance. The first two columns show pre- and post-ranking standard deviation of each portfolio. The next two columns present CAPM betas for each portfolio. The last columns present mean returns and CAPM alphas for post-ranking portfolios.

Portfolio	SD_t	SD_{t+1}	β_t	β_{t+1}	r_{t+1}	α_{t+1}	t-stat
1	2.50	3.25	0.50	0.65	0.86	0.15	[2.06]
2	3.69	4.08	0.81	0.87	0.97	0.16	[2.74]
3	4.58	4.76	1.01	1.01	0.93	0.05	[0.91]
4	5.34	5.26	1.17	1.10	1.03	0.08	[1.43]
5	6.12	5.89	1.30	1.21	1.11	0.09	[1.26]
6	6.96	6.52	1.43	1.30	1.05	−0.04	[−0.44]
7	7.98	7.36	1.56	1.38	0.90	−0.27	[−2.16]
8	9.12	8.06	1.67	1.43	0.69	−0.53	[−3.43]
9	10.42	8.69	1.71	1.41	0.31	−0.95	[−4.83]
10	13.07	9.20	1.68	1.29	−0.28	−1.55	[−6.09]