Toward Conditional Risk Parity: *Improving Risk Budgeting Techniques in Changing Economic Environments*

LIONEL MARTELLINI, VINCENT MILHAU, AND ANDREA TARELLI

LIONEL MARTELLINI
is a professor of finance at EDHEC Business School, scientific director of EDHEC-Risk Institute, and a senior scientific advisor at ERI Scientific Beta in Nice, France.
lionel.martellini@edhec.edu

VINCENT MILHAU
is the deputy scientific director of EDHEC-Risk Institute in Nice, France.
vincent.milhau@edhec.edu

ANDREA TARELLI
is a post-doctoral research fellow at Bocconi University in Milan, Italy.
andrea.tarelli@unibocconi.it

Risk parity has become an increasingly popular approach for portfolio construction within and across asset classes. In a nutshell, the goal of the methodology is to ensure that the constituent assets have the same contribution to the overall risk of the portfolio, as opposed to having the same dollar contribution, as they would in an equally weighted strategy. The notion of contribution to volatility was defined by Litterman [1996], and the difference between diversification in terms of dollars and diversification in terms of risk was pointed to by Qian [2005]. The first formal analysis of the "equal risk contribution" portfolio was subsequently given by Maillard et al. [2010], who establish its existence and uniqueness, derive a number of analytical properties, and propose numerical algorithms to compute the portfolio. An in-depth overview of the benefits and limits of risk parity portfolios is given in Roncalli [2013b].

While intuitively appealing and empirically attractive, risk parity (in short, RP) suffers from two major shortcomings. First, typical RP strategies used in an asset allocation context inevitably involve a substantial overweighting of bonds with respect to equities, which raises a serious concern in a low bond yield environment. Indeed, the presence of mean reversion in interest rates leads to an expected subsequent increase in bond yields, hence a poor future performance for bond portfolios (see Inker [2010] and Schachter and Thiagarajan [2011] for related arguments). Second, standard approaches to risk parity take volatility as a risk measure, which implies that upside risk is penalized as much as downside risk, in obvious contradiction to investors' preferences.

This article can be regarded as an attempt to address these two concerns via the introduction of a new class of risk parity strategies. We refer to them as "conditional risk parity" (CRP) strategies, in contrast to the standard "unconditional risk parity" (URP) strategies that are based upon historical volatility.

In a first step, we recognize that the volatility of a bond is related to its duration, and that duration is a decreasing function of the bond yield. As a result, it is possible to construct a volatility estimate such that the bond weight in the RP portfolio is increasing in the yield. In a second step, we follow Hallerbach [2003], Qian [2006], and Roncalli [2013b] in defining a new class of conditional RP portfolios with respect to downside risk measures such as semi-volatility, value-at-risk or expected shortfall. In this extension, the riskiness of an asset class is an explicit function of its expected return, and a lower expected return results in a higher risk, thus in a lower allocation in the risk parity portfolio.

The vast empirical research on stock and bond return predictability can be used at this stage to construct expected return estimators.

For instance, a large literature has documented the ability of various state variables, such as the dividend yield, to predict stock returns (see Fama and French [1988], Hodrick [1992], and Menzly et al. [2004], as well as Welch and Goyal [2008] for a survey and an evaluation of the predictive models over more than 100 years). Similarly, a simple formal argument borrowed from Campbell et al. [1997], as well as empirical evidence, suggests that the bond yield is a valuable predictor of bond returns. By relating the weights of risk parity portfolios to state variables that are observable on each date, these new approaches to risk parity reduce the reliance on historical estimates, which are by nature backward-looking, and better reflect current market conditions, including, notably, current yield levels.

Overall our results show that CRP portfolios constructed from downside risk measures are more responsive to yield changes than are their URP counterparts, and also that they involve a lower allocation to bonds in a low yield environment, precisely when downside risk in bonds is higher.

FROM UNCONDITIONAL TO CONDITIONAL RISK PARITY STRATEGIES

In this section, we recall the mathematical definition of risk parity portfolios, and we then define three forms of conditional risk parity (CRP) strategies.

Definition of Unconditional Risk Parity Strategies

We start from the formal definition of RP strategies laid out in Maillard et al. [2010], but we explicitly indicate the dependency with respect to time, because time variation in volatilities and expected returns is an essential aspect of our article. The investment universe consists of N assets, whose ex ante vector of expected returns and covariance matrix between dates t and $t+1$ are denoted with μ_t and Σ_t. Given a vector of weights $\boldsymbol{w}_t = (w_{1t}, \ldots, w_{Nt})'$, the ex ante portfolio volatility is thus $\sigma_{Pt} = \sqrt{\boldsymbol{w}_t' \boldsymbol{\Sigma}_t \boldsymbol{w}_t}$. Because it is a homogenous function of degree 1 in the weights, it admits the following decomposition (known as Euler decomposition):

$$\sigma_{Pt} = \sum_{i=1}^{N} c_{it}^{\mathrm{vol}}(\boldsymbol{w}_t), \quad c_{it}^{\mathrm{vol}}(\boldsymbol{w}_t) = w_{it}\frac{\partial \sigma_{Pt}}{\partial w_{it}} = w_{it} \times \left[\boldsymbol{\Sigma}_t \boldsymbol{w}_t\right]_i \tag{1}$$

The RP portfolio is defined by the condition that all contributions are equal and the budget constraint is satisfied:

$$c_{1t}^{\mathrm{vol}}(\boldsymbol{w}_t) = \cdots = c_{Nt}^{\mathrm{vol}}(\boldsymbol{w}_t) \text{ and } \boldsymbol{w}_t' 1 = 1 \tag{2}$$

As shown by Maillard et al. [2010], if all assets have the same pairwise correlations, the RP portfolio coincides with the "inverse volatility" portfolio, where weights are inversely proportional to volatility. This is the case, in particular, if there are only two assets (thus, only one correlation). In the more general situation of a non constant correlation, there is no known closed-form expression for the weights, and numerical techniques have to be employed to compute them.[1] As noted by Maillard et al. [2010], the RP portfolio represents a midpoint between the equally weighted and the minimum variance portfolios in that its volatility is comprised between those of these two portfolios. Risk parity appears to be a more meaningful approach to diversification than equal weighting because it does not ignore the differences between volatilities and correlations.

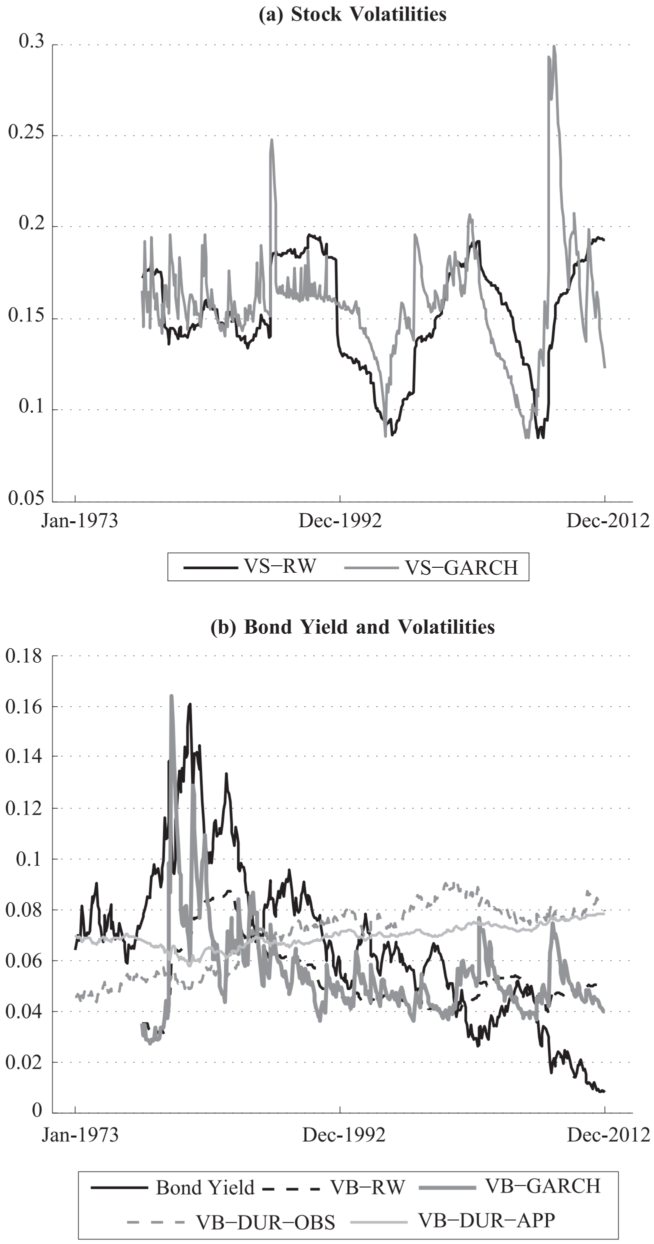
As appears from the definition, the only required input for the computation of RP weights is the covariance matrix. The traditional way of obtaining this parameter is to perform a rolling-window estimation, but this approach raises two statistical issues:

1. Sample risk: The estimate depends on a particular historical scenario, and is thus not fully representative of the true distribution of returns;
2. Stationarity risk: The rolling-window estimator approximates the *past* true covariance matrix, while the relevant parameter is the covariance matrix of future returns conditional on information available to date. If returns are not stationary, the two matrixes may be different.

The sample risk issue has been the focus of a large body of literature. To summarize it briefly, the sample covariance matrix is a reliable estimator of the true matrix as long as the sample size is much greater than the number of assets, and robustification techniques are needed when this condition is not satisfied.[2] In an asset allocation context, the number of constituents is typically small, so that robustification is not as crucially needed as in a portfolio construction exercise with several hundred

EXHIBIT 1
Estimated Bond and Stock Volatilities, January 1973–December 2012

All volatility measures are annualized. The signification of abbreviations is as follows: "VB-RW" and "VS-RW": rolling-window bond and volatilities; "VB-GARCH" and "VS-GARCH": GARCH(1,1) bond and stock volatilities; "VB-DUR-OBS": duration-based volatility computed with observed duration; "VB-DUR-APP": duration-based volatility computed with approximate duration.



constituents. On the other hand, stationarity risk is a concern because an asset that is a little risky in sample may have larger volatility out of sample, and thus contribute to risk more than expected.

The last problem with historical volatility is that it is not a good measure of downside risk. This is apparent in Exhibit 1, where we plot the rolling-window (in short, RW) volatility of a bond index (the Barclays U.S. Treasury). Although the volatility did not change much over the 1990–2012 period, the yield followed a decreasing trend to reach a floor close to zero, which leads us to expect that it will grow sooner or later. Thus, the downside risk of bonds has been increasing, and this increase was not reflected in the level of volatility.

In the next sections, we present several possible cures to these problems. Since the empirical application will be a stock-bond allocation exercise, we focus the discussion on the estimation of volatilities, which are the only inputs required to compute the RP portfolio.

CONDITIONAL RISK PARITY I: USING AN INSTANTANEOUS MEASURE OF VOLATILITY

In the first form of conditional risk parity, we still use volatility as a risk measure, but we attempt to relate it to current market conditions.

GARCH Volatility

The GARCH models introduced by Bollerslev [1986] deliver forward-looking volatility estimates by explicitly taking into account the dynamics of volatility.[3] The most simple version is a GARCH(1,1) where conditional expected returns on the stock and the bond are constant, and the conditional variance of the return depends on the lagged innovation to the asset (ARCH(1) term) and its own lagged value (GARCH(1) term). This method does not fully address the statistical concerns raised by RW volatility because even if it does not assume stationary returns as RW volatility does, the GARCH volatility is also subject to sample risk. Indeed, the parameters of the GARCH model have to be estimated from the data. Finally, it is not clear whether the GARCH volatility will do better than the RW one when it comes to reflecting the increasing downside risk in bonds in a period of falling yields.

Duration-Based Volatility

In this section, we argue that it is possible to construct a volatility estimator for the bond that does not rely on past returns and has a direct relationship to the yield level. The key element is the first-order approximation to a bond return as a function of the yield change. If C_t is the price of a coupon-paying bond, θ_t is its yield-to-redemption and D_t is its modified duration, then we have, by definition of the duration:

$$\frac{\Delta C_t}{C_t} \approx -D_t \times \Delta\theta_t \tag{3}$$

This suggests to estimate the volatility of the bond as the product of the duration and the estimated volatility of yield changes:

$$\hat{\sigma}_{DUR,t} = D_t \times \hat{\sigma}_{\theta,t} \tag{4}$$

For an index, which is a portfolio of coupon-paying bonds, the volatility estimator will take the same form, with the duration and the yield being computed as the weighted sums of the durations and yields of the constituents.

If it were not for the need to estimate yield volatility, this volatility estimator would be observable, as the duration is, and the sample risk issue would be completely addressed. In what follows, we assume for simplicity a constant yield volatility, so that bond volatility is proportional to the duration. As a result, the duration-based volatility will inherit an important property of duration, which is that it is decreasing in the yield.

More specifically, it can be shown that the modified duration of a coupon bond is a strictly decreasing function of the continuously compounded yield prior to the before-last cash-flow date, and is independent from the yield thereafter (since the modified duration of a zero-coupon bond equals its maturity).[4] An even more direct relationship between the duration and the yield can be obtained by using Campbell's approximation to duration (Campbell et al. [1997]). For a bond selling close to par, we have:

$$D_t \approx D_t^{\text{app def}} = \frac{1-e^{-m\theta_t}}{1-e^{-\theta_t}} \tag{5}$$

where m is the bond maturity (equality holds if the bond sells at par). A straightforward differentiation shows that the right-hand side is decreasing in the yield. In what follows, we will make a distinction between the two volatility estimators of the form (4) obtained with observed or approximate duration by referring to them as DUR-OBS and DUR-APP volatilities. Overall, these estimators, while not being fully observable, rely less on past returns than the RW volatility does, which alleviates statistical concerns, and they are decreasing in the yield. As will be empirically shown below, this helps them to capture the increasing downside risk of bonds.

CONDITIONAL RISK PARITY II: EXTENDING RISK PARITY TO DOWNSIDE RISK MEASURES

Even if historical volatility can be replaced by more forward-looking estimates, it remains true that volatility is a symmetric risk measure that does not make a distinction between downside risk and upside risk. Fortunately, the definition of an RP portfolio can be extended to other risk measures, provided they satisfy the condition of being homogenous of degree one in the weights. In details, if R_t is such a risk measure, the portfolio risk can be expressed as a sum of contributions from constituents:

$$R_t(w_t) = \sum_{i=1}^{N} c_{it}^R(\boldsymbol{w}_t), \quad c_{it}^R(\boldsymbol{w}_t) = w_{it}\frac{\partial R_t(\boldsymbol{w}_t)}{\partial w_{it}} \tag{6}$$

and the RP portfolio is defined by the condition $c_{1t}^R(\boldsymbol{w}_t) = \cdots = c_{Nt}^R(\boldsymbol{w}_t)$. In this section, we give several examples of such measures.

Gaussian Semi-Volatility and Value-at-Risk

To measure downside risk, a first idea is to replace volatility by the semi-volatility, which takes into account only negative returns:

$$\text{GSV}_t(\boldsymbol{w}_t) = \sqrt{\mathbb{E}_t\left[(\boldsymbol{w}_t'\ \boldsymbol{X}_{t,t+1})^2\ 1_{\{\boldsymbol{w}_t'\boldsymbol{X}_{t,t+1}\le 0\}}\right]} \tag{7}$$

where $\boldsymbol{X}_{t,t+1}$ is the vector of returns between dates t and $t+1$, and $\mathbb{E}_t$ denotes conditional expectation. In order to compute the partial derivatives of semi-volatility with respect to the weights, one needs an explicit expression for the portfolio semi-volatility. This can be computed in closed form under the assumption that the portfolio return is normally distributed. Thus, we define the Gaussian semi-volatility of the portfolio as the semi-volatility of a normally distributed variable with mean $\mu_{Pt} = \boldsymbol{w}_t'\boldsymbol{\mu}_t$ and variance $\sigma^2_{Pt} = \boldsymbol{w}_t'\Sigma_t\boldsymbol{w}_t$. Some algebra then shows that:

$$\mathrm{GSV}_t(\boldsymbol{w}_t) = [\sigma^2_{Pt} + \mu^2_{Pt}]\mathcal{N}\left(-\frac{\mu_{Pt}}{\sigma_{Pt}}\right) - \mu_{Pt}\sigma_{Pt} n\left(-\frac{\mu_{Pt}}{\sigma_{Pt}}\right) \quad (8)$$

where n and $\mathcal{N}$ are, respectively, the probability density function and the cumulative distribution function of the standard normal distribution. This expression allows for the derivation of contributions to risk, and the corresponding RP portfolio can be obtained with a numerical solver. An analytic differentiation of $\mathrm{GSV}_t(\boldsymbol{w}_t)$ with respect to μ_{Pt} also shows that the Gaussian semi-volatility is decreasing in the expected return. In other words, if two assets have the same volatility but different expected returns, the asset with the lower expected return will be deemed to be more risky. Hence, semi-volatility does capture the downside risk that is missed by volatility.

Another commonly used risk measure is the portfolio Value-at-Risk (VaR), which measures the size of "large" potential losses. Given a confidence threshold α, for which typical values are 95% or 99%, we define the portfolio VaR of order α as the quantile of order α of the distribution of the loss over one rebalancing period. Mathematically, this definition reads:

$$\mathbb{P}_t\left(-\boldsymbol{w}_t'\boldsymbol{X}_{t,t+1} \leq \mathrm{VaR}_t\right) = \alpha \quad (9)$$

where $\mathbb{P}_t$ denotes conditional probability. The VaR is clearly homogenous of degree one in the weights, so the Euler decomposition applies. Again, a distributional assumption is needed to compute the VaR and the risk contributions. The simplest option is the normality assumption, which leads to the following expression for the Gaussian VaR (see also Roncalli [2013a] for a study of the properties of this measure):

$$\mathrm{GVaR}_t(\boldsymbol{w}_t) = -\boldsymbol{w}_t'\mu_t + \mathcal{N}^{-1}(\alpha)\sqrt{\boldsymbol{w}_t'\Sigma_t\boldsymbol{w}_t} \quad (10)$$

Non-Gaussian Value-at-Risk

Because the normal distribution has thin tails, the normality assumption is hardly convenient to assess the size and the probability of potential losses in extreme events. The Cornish-Fisher correction (Cornish and Fisher [1938]) is often applied to incorporate the effects of higher order moments. It leads to the definition of the non-Gaussian VaR as:

$$\begin{aligned}\mathrm{NGVaR}_t(\boldsymbol{w}_t) = \mathrm{GVaR}_t(\boldsymbol{w}_t) + \Bigg[&\frac{q^2-1}{6} \\ &\times \mathrm{Sk}_t(\boldsymbol{w}_t) + \frac{q^3-3q}{24} \\ &\mathrm{Ku}_t(\boldsymbol{w}_t) - \frac{2q^3-5q}{36}\mathrm{Sk}_t(\boldsymbol{w}_t)^2 \; \sigma_{Pt}\Bigg] \quad (11)\end{aligned}$$

where $q = \mathcal{N}^{-1}(\alpha)$, and $\mathrm{Sk}_t(\boldsymbol{w}_t)$ and $\mathrm{Ku}_t(\boldsymbol{w}_t)$ are the portfolio skewness and (excess) kurtosis. Because these quantities depend on the co-skewness and the co-kurtosis matrix of the constituents, the use of this risk measure makes the estimation burden heavier.[5] Moreover, the derivation of risk contributions is complicated by the presence of the higher order moments. Nevertheless, it can still be carried out analytically (see Boudt et al. [2008] and the Appendix).

CONDITIONAL RISK PARITY III: DEVIATE FROM PARITY IN THE SHORT RUN WHILE BEING AT PARITY IN THE LONG RUN

A simple way of creating a dependence between weights and expected returns is to consider the maximum Sharpe ratio (MSR) portfolio, which by definition achieves the highest expected return per unit of risk taken. What is the relationship between the MSR portfolio and the concept of risk parity? First, the MSR portfolio achieves risk parity for a certain risk measure, which aggregates expected return and volatility. Indeed, it can be shown that it equates the contributions to the following measure:[6]

$$R_t(\boldsymbol{w}_t) = -\boldsymbol{w}_t'\tilde{\mu}_t + \lambda_{\mathrm{MSR},t}\sqrt{\boldsymbol{w}_t'\Sigma_t\boldsymbol{w}_t} \quad (12)$$

where $\tilde{\mu}_t$ is the vector of expected excess returns and $\lambda_{\mathrm{MSR},t}$ is the Sharpe ratio of the MSR portfolio, i.e., the maximum Sharpe ratio possible.

A second connection is established by Maillard et al. [2010]: the MSR equates the contributions to volatility if all assets have the same Sharpe ratio and the same pairwise correlations. In this case, the MSR weights are inversely proportional to volatilities, or, equivalently, to expected excess returns. In the case of two assets, namely stocks and bonds, this condition reduces to the equality between the two Sharpe ratios. The postulate that the two asset classes have the same long-term Sharpe ratio may not be strictly validated by the examination of a long historical period (see, e.g., the estimates of Dimson et al. [2008]), but it constitutes a reasonable agnostic working assumption in the absence of any prior information on the difference of performance between the two classes.

In the short run, however, this assumption is less convenient. For instance, after a large market downturn, equities may appear to be unusually cheap from a historical perspective, and thus have a more attractive Sharpe ratio than bonds do. As a conclusion, the MSR and the volatility-RP portfolios computed with long-term parameters will be equal, while the portfolios computed from parameters that depend on current market conditions will be different. The last form of CRP strategies that we introduce is the MSR portfolio computed with conditional estimates for volatilities and expected returns.

IMPLEMENTATION OF CRP STRATEGIES

We now turn to the practical implementation of the various forms of CRP strategies in a stock-bond allocation context. We use monthly data over the period 1973–2012. The stock is represented by the CRSP value-weighted index of the S&P 500 universe and the bond by the Barclays U.S. Treasury index. Both series incorporate the re-investment of dividends or coupons. We also download the time series of bond index duration and yield-to-redemption. In this section, we explain how the various volatility and expected returns are estimated and we study the relationship between the bond weight and the yield level for each RP strategy.

URP and CRP I Strategies

In these two strategies, the weight of each asset class is inversely proportional to its volatility, so that we have to estimate stock and bond volatilities.

Estimating stock and bond volatilities. For the URP strategy, we follow the widespread practice of taking a historical volatility estimate. Volatility is estimated from monthly returns over a 60-month rolling window. The first form of CRP I strategies (CRP-VOL-GARCH) relies on GARCH volatilities, which we extract from a GARCH(1,1) model. In order to avoid look-ahead bias, we repeat the estimation over a 60-month rolling window. The other two forms of CRP I strategies (CRP-VOL-DUROBS and CRP-VOL-DURAPP) use duration-based estimates for the bond volatility, stock volatility being estimated as the RW volatility. The two strategies differ by the duration measure: in CRP-VOL-DUROBS, it is the observed duration of the bond index, and in CRP-VOL-DURAPP, it is the approximate duration computed through (5). In this equation, the parameter m is chosen in such a way as to equate the historical averages of the observed and approximate durations; this criterion leads to $m = 5.37$ years. Finally, in both CRP-VOL-DUROBS and CRP-VOL-DURAPP, the yield volatility is estimated as the sample volatility of monthly yield changes over the 1973–2012 period. This volatility is 1.48% in annual terms.

The time series of volatilities shown in Exhibit 1 and the descriptive statistics reported in Exhibit 2 point to the existence of two groups of bond volatility measures. First, the RW and GARCH volatilities have close average values, of 5.3% and 5.4% respectively, and a strong correlation of 60.7%. The picture also shows that their historical patterns are similar. After a high regime, above 6%, between April 1980 and October 1983, and a peak at 8.72% in 1984 for the RW estimate, both volatilities went down to lower levels for a long period.

Second, the two duration-based volatility estimates have higher average values, of 7.2% for the DUR-OBS one and 7.0% for the DUR-APP one; this indicates that these two measures assign a higher risk to the bond than the RW and the GARCH volatilities do. They have followed a different pattern in the sample, starting from relatively low levels, below 7%, at the beginning of the sample, and displaying an upward trend from 1984 onward. This movement is explained by the decreasing trend of the yield, and it results in negative correlations between the duration-based volatilities and the volatilities in the first group. For the stock, the RW and GARCH estimates also appear to be close to each other, both in average and correlation, although the latter tends to exacerbate volatility peaks.

Weights of CRP I strategies. Panels (a) to (d) in Exhibit 3 show the weights of the URP strategy and the three variants of the CRP I strategy. (We will comment on the other panels in the next sections.) A common

EXHIBIT 2

Statistics on Stock and Bond Volatilities, January 1978–December 2012

The signification of abbreviations is as follows: "VB-RW" and "VS-RW": rolling-window bond and volatilities; "VB-GARCH" and "VS-GARCH": GARCH(1,1) bond and stock volatilities; "VB-DUR-OBS": duration-based volatility computed with observed duration; "VB-DUR-APP": duration-based volatility computed with approximate duration. In Panel (b), diagonal elements are volatilities and off-diagonal elements are correlations.

(a) Averages of Volatility Measures

VB-RW	VB-GARCH	VB-DUR-OBS	VB-DUR-APP
0.053	0.054	0.072	0.070

VS-RW	VS-GARCH
0.152	0.158

(b) Correlations and Volatilities of Volatility Measures

	VB-RW	VB-GARCH	VB-DUR-OBS	VB-DUR-APP
VB-RW	0.013			
VB-GARCH	0.607	0.018		
VB-DUR-OBS	−0.555	−0.432	0.011	
VB-DUR-APP	−0.566	−0.424	0.836	0.005

	VS-RW	VS-GARCH
VS-RW	0.029	
VS-GARCH	0.405	0.032

point of all the strategies is that they contain, on average, more bonds than stocks, which is explained by the fact that the bond is in general less volatile than the stock. Nevertheless, the bond is more predominant in URP and CRP-VOL-GARCH portfolios; for instance, at the end of the sample (on December 31, 2012), the bond weights are respectively 79.8% and 82.9% in these two portfolios, while the strategies CRP-VOL-DUROBS and CRP-VOL-DURAPP contain only 70.9% and 71.2% of bonds. Overall, it appears that the replacement of the RW volatility by the GARCH one has no substantial impact on the stock-bond allocation, while the use of a duration-based volatility estimate leads to a decrease in the bond weight in a period of falling interest rates. Thus, the two duration-based measures capture, at least in part, the increasing downside risk of the bond that is associated with the decreasing yield.

To assess the difference in the behavior of the strategies with respect to bond yield, one can look at the correlation between the bond weight and the yield. A high correlation means that the bond weight tends to decrease when the yield decreases. To the extent that the yield has the ability to predict future bond returns (which we shall confirm below), this implies that the strategy tends to allocate less to the bond when it gets more expensive, which is certainly a desirable property.

In order to go beyond the statistical link between the two quantities, one can also compute the "concordance rate," defined as the percentage of months where the weight and the yield moved in the same direction. If the rate is greater than 50%, then, in the majority of cases, the strategy makes the "right" decision about bond weighting: it reduces the allocation if the yield goes down, and increases it if the yield goes up. As appears from the first five rows of Exhibit 4, the strategies URP and CRP-VOL-GARCH tend to do the opposite, allocating more to bonds as the yield decreases. On the other hand, the behavior of strategies CRP-VOL-DUROBS and CRP-VOL-DURAPP is more in line with the variation of the yields, as can be seen from the positive correlations and the concordance rates greater than 50%.

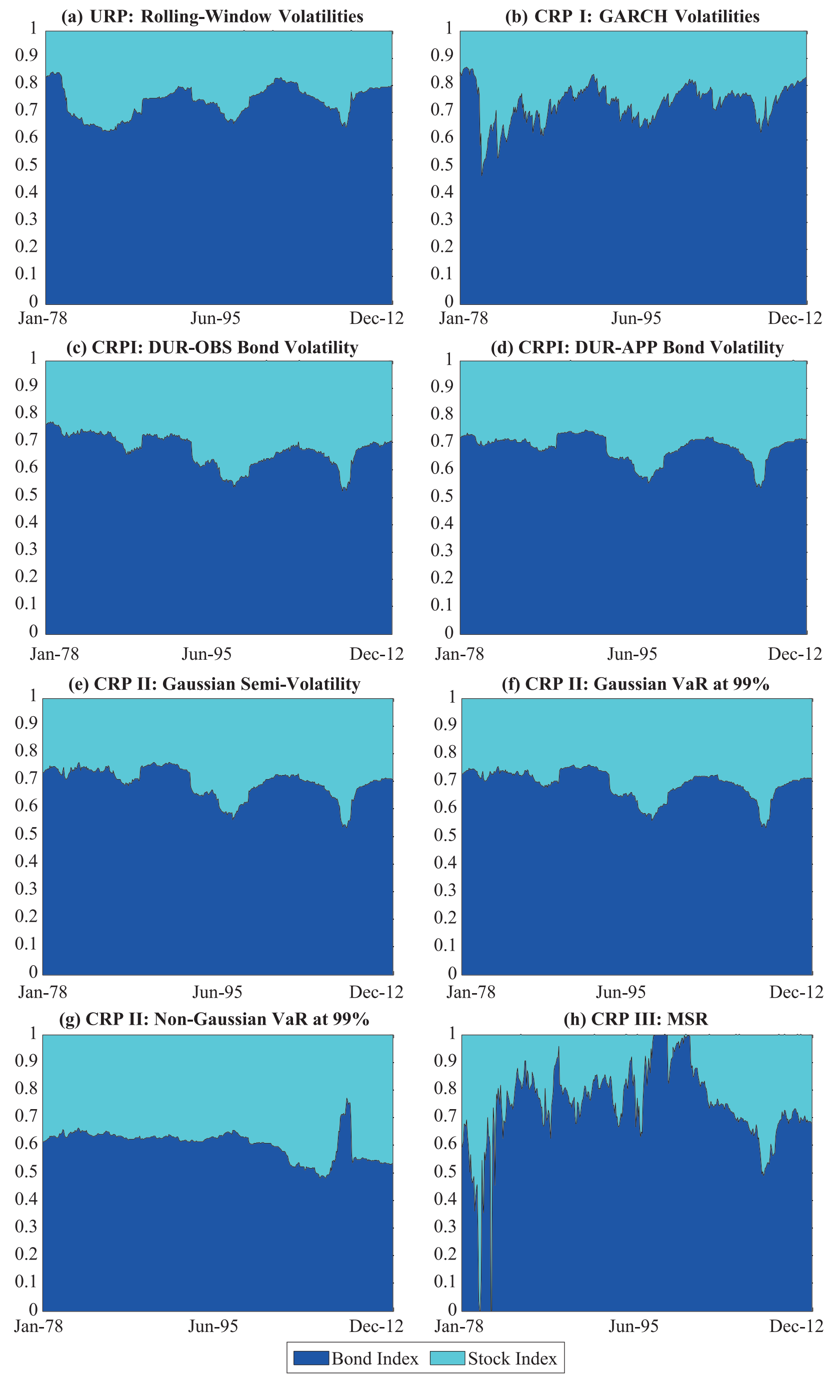
CRP II Strategies

The main difference between CRP II and CRP I strategies is that the former require the estimation of bond and stock expected returns in addition to volatilities.

Exhibit 3

Weights of URP and CRP Strategies, January 1978–December 2012

This figure shows the weights of bond-stock risk parity strategies. Strategies URP and CRP I use volatility as a risk measure; strategies CRP II use a downside risk measure; strategy CRP III is the maximum Sharpe ratio portfolio. The DUR-OBS and the DUR-APP bond volatility estimates are proportional respectively to the observed duration and the approximate duration (5). In strategies CRP II and CRP III, the expected returns on the stock and the bond are obtained by shrinking a forecast obtained from a predictive regression towards a prior.



Estimating bond and stock expected returns. Since the sample mean of past returns is known to be a very inaccurate estimator of expected returns (Merton [1980]), we construct alternative estimators by using observable predictors.

Predicting stock returns. We follow Fama and French [1988] in regressing the stock return on the dividend-price (DP) ratio, computed as the sum of dividends paid in the previous 12 months, divided by the current price index.[7] If h denotes the prediction horizon, expressed as a number of months, the regression equation reads:

$$\ln \frac{S_{t+h} + \sum_{k=1}^{h} d_{t+k}}{S_t} = \alpha_S + \beta_S \mathrm{DP}_t + \varepsilon_{S,t+h} \tag{13}$$

where S_t and DP_t are the price index and the DP ratio at the end of month t, and d_{t+k} is the dividend paid in month $t+k$. The "forecasted" return is then $\hat{\mu}_{St}^{\text{forecast}} = \hat{\alpha}_S + \hat{\beta}_S \mathrm{DP}_t$. We take the prediction horizon to be $h = 12$ months, which gives an R^2 of 9.2% and a slope coefficient of 3.983 with a t-statistic of 3.049 (including the autocorrelation adjustment of Newey and West [1987]).[8]

The R^2 of 9.2% indicates that a large fraction of the variance is not explained by the predictor. Moreover, the coefficients α_S and β_S are imperfectly estimated. Overall, there remains a substantial uncertainty over the value of the true expected return, which suggests that it would be safer not to rely entirely on the forecasted return. To mitigate the sample noise, we follow the well-established practice of shrinking the estimator toward a prior (see Barry [1974], Jorion [1985] and DeMiguel et al. [2009]). This prior is constructed as

$$\hat{\mu}_{St}^{\text{prior}} = r_t + \bar{\lambda}_S \hat{\sigma}_{S-RW,t} \tag{14}$$

where r_t is the short-term risk-free rate—represented by the three-month Treasury bill secondary market rate from the Fed website, $\hat{\sigma}_{S-RW,t}$ is the RW volatility and $\bar{\lambda}_S$ is the long-term Sharpe ratio of U.S. stocks. The figures of Ibbotson [2013] imply a value of 0.41 for $\bar{\lambda}_S$. The posterior estimator for expected return is eventually obtained as a combination of the forecast and the prior:

$$\hat{\mu}_{St}^{\text{posterior}} = \omega_S \hat{\mu}_{St}^{\text{prior}} + (1-\omega_S)\hat{\mu}_{St}^{\text{forecast}} \tag{15}$$

To compute the shrinkage intensity, we follow Vasicek [1973] in taking $\omega_S = S_S^{\text{forecast}}/(S_S^{\text{prior}} + S_S^{\text{forecast}})$, where S_S^{forecast} and S_S^{prior} are the variances of the two estimators. The variance of the forecast in a month t can be obtained from the covariance matrix of the estimators $(\hat{\alpha}_S, \hat{\beta}_S)$, and S_S^{forecast} is then estimated as the average of the variances across dates. For the prior, we assume that the uncertainty over the values of volatilities is negligible compared to the uncertainty over the value of the unconditional expected excess return. This amounts to taking S_S^{prior} equal to the estimated variance of the long-term average excess return.

Because Ibbotson's estimate is a sample mean, we estimate its variance by dividing the long-term variance of excess returns, $\overline{\sigma_S}^2 = 0.202^2$, by the number of years contained in his sample, $n_{\text{obs}} = 87$. Eventually, we obtain a shrinkage intensity $\omega_S = 0.57$.

Predicting bond returns. Fama [1984], Fama and Bliss [1987], and Cochrane and Piazzesi [2005] found that forward rates have the ability to predict excess returns on long-term bonds, and several articles have proposed multivariate prediction models for the prediction of Treasury bond returns (see, e.g., Ilmanen [1995 and 1997], Baker et al. [2003], Duffee [2011], and Joslin et al. [2014]).

But as we did for the stock, we favor a univariate approach here. The bond yield is a natural candidate because the bond price is a monotonic (decreasing) function of the yield, and the yield, as interest rates, has a tendency to revert to a long-term mean.[9] Hence, low yields are likely to be followed by high yields, that is, by low bond prices, while high yields are advanced indicators of high future bond prices. Thus, the yield is a natural predictor of bond returns in the sense that it is a "cheapness" indicator of the bond relative to historical standards.

A formal justification for the use of this variable can also be given, based on the approximation of Campbell et al. [1997], to the one-period log return on a coupon-paying bond. If the bond sells close to par, we have:

$$\ln \frac{C_{t+1}}{C_t} \approx -(D_t - 1)(\theta_{t+1} - \theta_t) + \theta_t \tag{16}$$

This decomposition involves the usual income term (θ_t) and contribution from yield change (first term). Taking expectations in both sides and assuming, as Fama and Bliss do, that the yield is close to a random walk—so that the expected yield change is close to a constant—we obtain that the expected bond return is approximately an affine function of the yield. This analysis motivates the construction of a forecast for bond returns as:

$$\hat{\mu}_{Bt}^{\text{forecast}} = \hat{\alpha}_B + \hat{\beta}_B \theta_t \tag{17}$$

where $\hat{\alpha}_B$ and $\hat{\beta}_B$ are the estimated intercept and slope coefficient in a regression of the total future bond return (i.e., with coupons re-invested) on the current yield. At the one-year prediction horizon, the quality of the fit is better than for the stock, with an R^2 of 31.7%. We find a positive slope coefficient of 1.032, which is statistically significant even after adjusting for autocorrelation in residuals. This confirms that the yield has forecasting ability.

As for the stock, we shrink the forecast toward a prior constructed to imply a Sharpe ratio of 0.26, which is the value implied by the long-term figures of Ibbotson [2013]. The shrinkage intensity is fixed according to the same rule as for the stock, which results in $\omega_B = 0.36$. Thus, more confidence is given to the forecast than in the case of the stock.

Weights of RP strategies. Unlike for volatility, the weights of the RP strategies defined with respect to downside risk measures are not simply proportional to the reciprocals of the constituents' risk measures. They must be computed by numerically solving the system $c_{1t}^R(\boldsymbol{w}_t) = \cdots = c_{Nt}^R(\boldsymbol{w}_t)$, where the c_{it}^R are the risk contributions. The Appendix gives analytical expressions for the contributions to the three downside risk measures. They require estimates for the volatilities and the expected returns of the stock and the bond. We use the RW volatility for the stock and the DUR-APP volatility for the bond, and the shrinkage estimators for expected returns. But the risk contributions also involve a new set of parameters. First, they depend on the stock-bond correlation; we estimate it over a 60-month rolling window, as the stock volatility. Second, the non-Gaussian VaR depends on the co-skewness and the co-kurtosis matrix of the stock and the bond, denoted as $\boldsymbol{M}_{3t}$ and $\boldsymbol{M}_{4t}$. In order to avoid the estimation of odd-order moments, we take $\boldsymbol{M}_{3t}$ equal to zero.

In the estimation of $\boldsymbol{M}_{4t}$, we follow the idea of relating bond variance to duration. In line with approximation (3), we estimate the elements of the co-kurtosis matrix as:

$$\mathbb{E}_t\left[\tilde{r}_{S,t,t+1}^i \tilde{r}_{B,t,t+1}^j\right] \approx (-D_t^{\text{app}})^j \, \mathbb{E}_t\left[\tilde{r}_{S,t,t+1}^i (\Delta\theta_t - \mu_{\theta t})^j\right] \tag{18}$$

for any choice of indexes such that $i + j = 4$. In this equation, $\tilde{r}_{S,t,t+1}$ and $\tilde{r}_{B,t,t+1}$ denote the centered simple returns on the stock and the bond, and $\mu_{\theta t}$ is the mean of yield changes. We are thus back to the estimation of the co-kurtosis matrix of the stock and yield changes.

Having estimated the higher order comoments, we can proceed with the calculation of risk contributions as shown in the Appendix and the numerical solution of the system $c_{1t}^R(\boldsymbol{w}_t) = \cdots = c_{Nt}^R(\boldsymbol{w}_t)$. Throughout these computations, we take the confidence levels of the two VaR measures to be $\alpha = 99\%$. Panels (e) to (g) in Exhibit 3 show the weights of the three RP portfolios. The weights of the two strategies that rely on Gaussian risk measures (semi-volatility or VaR) are visually very close to each other and to those of the CRP-VOL-DUR-APP portfolio (Panel (d) in Exhibit 3). It is the non-Gaussian VaR that stands out among the three downside risk measures, with a more balanced allocation between stocks and bonds, the average bond weight being close to 55%.

Even if the weights are not inversely proportional to the constituents' risk measures, this observation suggests that with the non-Gaussian VaR, the stock and the bond have similar levels of riskiness. The time series average of the ratio of stock risk to bond risk is a simple indicator that helps to verify this fact: it is 2.18 with volatility, 2.44 with the Gaussian semi-volatility, 2.32 with the Gaussian VaR, and only 1.38 with the non-Gaussian one. Nevertheless, the allocation with the non-Gaussian VaR is different from a fixed-mix 50%/50% portfolio, since the bond weight has a tendency to decrease over time. This can be related to the decrease in the yield, which increases the DUR-APP volatility and decreases the expected return, two effects that contribute to increase the VaR.

Exhibit 4 shows the concordance indicators for the three CRP II strategies. With the two Gaussian measures, the correlation between the bond weight and the yield is not significantly increased with respect to the case where volatility is taken as a risk measure. It is comprised between 40% and 50%, when a correlation of 48% could be achieved by using the DUR-OBS volatility for the bond. A slightly higher improvement is achieved in the concordance rates: in more than 71% of months, the yield and the bond weight moved in the same direction, while the rate was at best 67.3% with strategies relying on volatility. The largest increases in correlation and concordance rate are obtained by choosing the non-Gaussian VaR as a risk measure.

Exhibit 4

Correlations and Concordance Rates for Weights of RP Strategies Invested in Stocks and Bonds, January 1978–December 2012

The left column shows the historical correlations between the bond weight and the bond yield. The right column contains the concordance rates, defined as the percentages of months in which the bond weight and the yield moved in the same direction. The strategies are the unconditional risk parity strategy (URP), the two conditional risk parity (CRP I) strategies based on volatility, the three CRP II strategies based on downside measures, and the CRP III strategy, which is the maximum Sharpe ratio portfolio.

	Corr. (w, Bond Yield)	Concordant Months (%)
URP	−0.476	45.346
CRP I (GARCH volatilities)	−0.390	48.687
CRP I (DUR-OBS bond volatility)	0.480	62.768
CRP I (DUR-APP bond volatility)	0.261	67.303
CRP II (Gaussian semi-volatility)	0.465	73.866
CRP II (Gaussian VaR at 99%)	0.400	71.957
CRP II (non-Gaussian VaR at 99%)	0.611	74.582
CRP III (MSR)	−0.051	71.241

CRP III Strategy

The CRP III portfolio is the MSR portfolio of the stock and the bond. It is known that its weights are extremely sensitive to expected return estimates (Best and Grauer [1991]), and that the sample mean, while being a convergent and unbiased estimator of the true expected return, has too large a variance at usual sample sizes to be a reliable estimate (Merton [1980]). Hence, a robust estimation of expected returns is of special importance here. We thus again use shrinkage estimators, but we modify the priors in such a way that the portfolio is at risk parity in the long run. In CRP II strategies, the priors imply constant Sharpe ratios equal to the long-term values (0.26 for the bond and 0.41 for the stock). For the CRP III strategy, we construct the priors by assuming identical Sharpe ratios for both assets; we take this common value to be the long-term Sharpe ratio of the bond, namely 0.26.[10] Expected excess returns are then estimated by subtracting the short-term risk-free rate.

Panel (h) in Exhibit 3 shows that in spite of this shrinkage, which should mitigate the effects of estimation errors, the weights of the MSR portfolio exhibit a much larger variability over time than those of the URP, CRP I, and CRP II portfolios. Hence, it is the MSR portfolio that will have the largest turnover. This result confirms the high sensitivity of MSR weights to time variation in input parameters, in particular expected returns.

Exhibit 4 takes a closer look at the link between the bond weight and the yield in the MSR portfolio. The concordance rate of 73.4%, is as high as with the CRP strategies relying on downside risk measures, while the correlation is negative. The former indicator suggests that the relationship between the two variables is rather reliable, but the latter points to a non linear dependency.

BENEFITS OF CRP STRATEGIES IN PERIODS OF INCREASES IN INTEREST RATES

The previous sections have shown that the various forms of CRP strategies tend to have a lower bond allocation than standard URP strategies in low interest rate environments. It remains to assess the impact of this feature on the performance of the RP portfolio in the scenario of a mean-reversion of bond yields back up to their historical mean levels. To this end, we simulate an increase in interest rates starting from historically low levels, and compare the performances of the various forms of RP strategies.

Generating the Stochastic Scenarios

We simulate monthly scenarios for the following processes: bond and stock prices, dividend-price (D/P) ratio, yield-to-redemption, and duration. Since the assumption that interest rates are mean-reverting is critically important for the analysis conducted in this section, we model the yield-to-redemption of the bond index as an AR(1) process (we denote simulated variables with tildas):

EXHIBIT 5
Parameter Values in Monte-Carlo Simulation

The risk premium and the volatility of the stock, $x_{S,t}$ and $\sigma_{S,t}$, are constant in the first scenario and vary across years 1 to 5 in the second scenario. Scenario 1 corresponds to a relatively slow increase in interest rates with no impact on the equity market, and scenario 2 corresponds to a rapid increase with a strong reaction of the equity market in years 2 and 3. For each mean-reverting process (bond yield, short-term rate, dividend-price ratio), α^1, α^0 and σ denote respectively the slope coefficient, the intercept and the volatility of innovations in the AR(1) model. $\tilde{\theta}_0, \tilde{r}_0$ and $\widetilde{DP}_0$ are the initial values of the processes.

Stock Price

	Scenario 1					Scenario 2				
	Y1	Y2	Y3	Y4	Y5	Y1	Y2	Y3	Y4	Y5
$x_{S,t}$	0.06	0.06	0.06	0.06	0.06	0.06	0.00	−0.10	0.05	0.10
$\sigma_{S,t}$	0.15	0.15	0.15	0.15	0.15	0.15	0.20	0.25	0.25	0.20

Bond Yield

	Scenario 1	Scenario 2
α^1_θ	0.9638	0.9120
α^0_θ	0.0024	0.0058
σ_θ	0.0042	0.0042
$\tilde{\theta}_0$	0.0086	0.0086

Short-Term Rate

	Scenario 1	Scenario 2
α^1_r	0.9638	0.9120
α^0_r	0.0019	0.0047
σ_r	0.0052	0.0051
$\tilde{r}_0$	0.0005	0.0005

Dividend-Price Ratio / **Correlations**

Dividend-Price Ratio		Correlations			
α^1_{DP}	0.00026	$\rho_{S\theta}$	−0.1158	ρ_{SDP}	0.9915
α^0_{DP}	−0.8833	$\rho_{\theta DP}$	0.1889	ρ_{Sr}	−0.0122
σ_{DP}	0.0017	$\rho_{\theta r}$	0.7241	$\rho_{DP,r}$	0.0457
$\widetilde{DP}_0$	0.0227				

$$\tilde{\theta}_{t+1} = \alpha^0_\theta + \alpha^1_\theta \tilde{\theta}_t + \sigma_\theta \eta_{\theta,t+1} \tag{19}$$

The dividend-price ratio and the short-term interest rate are also assumed to follow mean-reverting processes, and we denote their parameters with the same symbols by using subscripts DP and r.

To simulate bond duration and returns that are consistent with the simulated yield, we use the approximated duration (5) with $m = 5.37$ years, and the monthly log bond return evolves as:

$$\log(1+\widetilde{X}_{B,t,t+1}) = \widetilde{D}_t^{\text{app}}(\tilde{\theta}_{t+1} - \tilde{\theta}_t) + h\tilde{\theta}_t \tag{20}$$

where $h = \frac{1}{12}$. As far as the stock index is concerned, we take the expected excess return to be a piecewise constant function of time (it will be specified later), so the dynamics of the log stock return read:

$$\log(1+\widetilde{X}_{S,t,t+1}) = h[\tilde{r}_t + x_{St}] + \sqrt{h}\sigma_S \eta_{S,t+1} \tag{21}$$

We take December 31, 2012, to be the starting date of our simulations, so the initial values for the bond yield, the short-term rate and the dividend-price ratio are the values on this date. The values of the parameters that govern the dynamics of the processes are given in Exhibit 5. The parameters of the DP are fixed in such a way as to match the historical average, autocorrelation and volatility, and the correlations between innovations are set to their historical values.

To fix the other parameters, we consider two economic scenarios, with a more or less fast increase in the level of interest rates in the first five years. In the first one, rates grow progressively, without a significant impact on the equity market. The chosen autoregression coefficient, α^1_θ, implies that it takes an average of five years for the yield to grow from an initial level of 0.86% to 6%; the 6% corresponds roughly to 90% of the historical mean over the period 1973–2012, which is 6.63%. In this scenario, the equity risk premium and volatility are constant across years 1 to 5. In the second scenario, interest rates increase at a much faster pace. It takes an average of two years for the yield to grow from its initial level to 6%, which implies a lower autoregression coefficient.[11] In this scenario, the central bank fails at controlling inflation expectations, which results in a bear equity market in years 2 and 3, with low expected returns and high volatility, before a recovery in years 4 and 5. In each scenario, the speed of mean reversion of the short-term rate is taken equal to that of the bond yield, and the constant terms and the volatility of innovations are chosen so as to match the historical means and volatilities.

PERFORMANCE OF RISK PARITY STRATEGIES

We now turn to the comparative analysis of several RP strategies: the URP strategy, the CRP I strategy relying on duration-based volatility for the bond, one CRP II strategy (the one relying on the non-Gaussian VaR as a risk measure, because it implies the best historical concordance indicators in Exhibit 4), and the CRP III strategy. We do not report the results for GARCH volatilities, since the behavior of the strategy has been shown to be close to that of URP.

We also introduce a fixed-mix benchmark, whose weights are the average weights over the 1973–2012 period of the URP strategy. The resulting allocation is 25.96% to the stock and 74.04% to the bond. Each of the RP strategies is implemented according to the same rules as in the historical period: for URP, bond and stock volatilities are estimated over a 60-month rolling window, and expected returns are estimated as linear combinations of a forecast and a prior.

The forecasts for the stock and the bond are affine functions of the DP ratio and the yield respectively, the coefficients of the functions being taken equal to the coefficients estimated over the 1973–2012 window. The priors are based on estimates for long-term Sharpe ratios; in the CRP II portfolio, the long-term values are respectively 0.26 and 0.41 for the bond and the stock, and they are both equal to 0.26 in the CRP III portfolio. Note that in any case, the expected returns are ex ante different from the true values assumed in the data-generating process. Hence, our results capture the effects of estimation errors on the various strategies.

Exhibit 6 reports simulated performance statistics for years 1 to 5 in the simulation. For each year, these statistics are first computed in the time series, and then averaged across 2,000 simulated paths. Because the yield starts from 0.05% and has a long-term mean of 6.63%, it has a tendency to grow in each scenario. In Scenario 2, where the speed of mean reversion is larger, reversion to the long-term mean is almost complete after five years, with an expected yield of 6.6%.

In parallel with this interest rate growth, the bond experiences negative returns. In scenario 1, where the yield goes up at a relatively slow rate, excess returns are strongly negative in the first two years (e.g., −9.7% in year 1), and they stay subsequently negative until year 5. In scenario 2, losses are much more severe in the first years, with the bond index underperforming the cash account by 18.4% in 2013. In the last two years, the yield keeps increasing, but at a slow rate, because it is already close to its stationary value. The income term in (16) is then sufficient to compensate the negative contribution of the yield change.

In both scenarios, all portfolios have negative returns in year 1, due to the negative bond returns. But as Exhibit 3 highlights, it is the CRP II strategy relying on non-Gaussian VaR that has the lowest bond allocation (53.3%) at the end of 2012, which is the beginning of the simulations. Rather unsurprisingly, it is also this portfolio that suffers the least from the poor performance of bonds in years 1 and 2. However, in scenario 2, it has the worst performance in year 3, underperforming the fixed-mix and the URP strategies by more than 1%. Indeed, this year is characterized by a strong equity bear market, so that the higher stock weight penalizes the portfolio performance.

Year 3 also appears as a bad year for the CRP II portfolio from a risk standpoint, because it inherits the large volatility of stocks; the portfolio volatility in this year is 12.3%. But in other years, the CRP II strategy is the best performer. In scenario 1, things are quieter on the equity side. There is no bear market, and CRP II always outperforms the other portfolios.

The MSR portfolio does not show particularly attractive performances in either of the two scenarios. In fact, it is the worst performer in all years, except for year 3 in the second scenario. While this poor score cannot be attributed with certainty to the effect of estimation errors in expected returns, it seems likely that this parameter plays a role. Indeed, in the first simulated year, the MSR underperforms the fixed-mix and the URP strategies, while Exhibit 3 shows that it has a lower bond weight than these two portfolios at the end of 2012 (68.1% versus 74.04% and 79.8%, respectively). Hence, it is unlikely that the bad performance comes from an overweighting of bonds.

Overall, these numbers illustrate in a quantitative manner the fact that overweighting bonds after a long period of decrease in interest rates, as a typical risk parity approach based on historical volatility does, leads to strongly negative returns when interest rates start to increase back to higher levels, especially if this increase is rapid. This problem can be at least partially addressed by using a CRP strategy, and the one based on non-Gaussian VaR shows interesting performances from this perspective. Moreover, the results suggest that the imperfection in expected return estimates, which impacts CRP II strategies, is more detrimental to the performance of the MSR portfolio.

CONCLUSION

The traditional approach to constructing risk parity portfolios is to equate the contributions of constitu-

Exhibit 6

Simulated Performance Statistics of Risk Parity and Fixed-Mix Strategies Invested in Stocks and Bonds

In Panels (a) and (c), the first two lines are the annual averages of the bond yield and the annual excess returns over the risk-free rate of the bond, the stock, a fixed-mix portfolio, and four implementations of the risk parity strategy (URP, CRP I, CRP II, and CRP III). Panel (b) and (d) report annual volatilities. The statistics are first computed along each of the 2,000 simulated paths, and are then averaged across paths.

(a) Scenario 1: Average Interest Rate and Excess Returns

	Y1	Y2	Y3	Y4	Y5
Yield (av.)	0.020	0.036	0.047	0.054	0.059
Bond index	−0.097	−0.053	−0.034	−0.015	−0.005
Stock index	0.061	0.062	0.054	0.059	0.061
Fixed-mix	−0.053	−0.020	−0.008	0.007	0.015
URP	−0.061	−0.021	−0.006	0.011	0.020
CRP I (DUR-APP bond volatility)	−0.079	−0.037	−0.021	−0.004	0.005
CRP II (non-Gaussian VaR at 99%)	−0.018	0.009	0.014	0.025	0.031
CRP III (MSR)	−0.084	−0.039	−0.023	−0.005	0.003

(b) Scenario 1: Volatilities

	Y1	Y2	Y3	Y4	Y5
Yield	0.015	0.015	0.015	0.015	0.014
Bond index	0.075	0.072	0.071	0.069	0.068
Stock index	0.146	0.147	0.148	0.146	0.147
Fixed-mix	0.071	0.070	0.069	0.068	0.067
URP	0.070	0.070	0.070	0.070	0.071
CRP I (DUR-APP bond volatility)	0.071	0.068	0.067	0.066	0.065
CRP II (non-Gaussian VaR at 99%	0.084	0.087	0.086	0.084	0.083
CRP III (MSR)	0.072	0.071	0.071	0.071	0.071

(c) Scenario 2: Average Interest Rate and Excess Returns

	Y1	Y2	Y3	Y4	Y5
Yield (av.)	0.031	0.055	0.063	0.065	0.066
Bond index	−0.184	−0.047	−0.010	0.007	0.010
Stock index	0.061	0.007	−0.102	0.036	0.102
Fixed-mix	−0.117	−0.029	−0.028	0.020	0.038
URP	−0.127	−0.029	−0.029	0.021	0.039
CRP I (DUR-APP bond volatility)	−0.157	−0.039	−0.017	0.012	0.021
CRP II (non-Gaussian VaR at 99%)	−0.064	−0.015	−0.045	0.027	0.054
CRP III (MSR)	−0.166	−0.042	−0.013	0.009	0.013

(d) Scenario 2: Volatilities

	Y1	Y2	Y3	Y4	Y5
Yield	0.015	0.015	0.015	0.015	0.015
Bond index	0.077	0.070	0.069	0.068	0.068
Stock index	0.146	0.193	0.243	0.244	0.196
Fixed-mix	0.073	0.077	0.085	0.086	0.076
URP	0.072	0.078	0.087	0.087	0.076
CRP I (DUR-APP bond volatility)	0.073	0.069	0.070	0.069	0.066
CRP II (non-Gaussian VaR at 99%)	0.085	0.105	0.123	0.119	0.095
CRP III (MSR)	0.074	0.069	0.069	0.068	0.067

ents to portfolio volatility while using rolling-window estimates as volatility measures. This approach raises a number of concerns, because such estimates rely heavily on a specific sample and do not necessarily provide a good assessment of current market conditions. More generally, the adoption of volatility does not capture dissymmetric investors' preferences for upside versus downside uncertainty.

We argue that the use of an alternative bond volatility estimate, which is proportional to the duration, alleviates the reliance on the historical sample, and we show that this volatility measure better reflects the increased downside risk associated with a low-yield environment. The latter effect can be reinforced by replacing volatility with a downside risk measure, such as semi-volatility or VaR, especially if this measure takes into account the non-Gaussian nature of return distributions.

Our empirical results show that "conditional risk parity" strategies, that is, strategies that use the duration-based volatility or a dissymmetric risk measure, are indeed more responsive to yield changes than a classical form of risk parity portfolio based on historical volatilities. This new generation of risk parity strategies introduces a new dimension, that is directional risk, as reflected in particular in the first moment (expected return) of asset return distributions.

By reintroducing expected returns into the risk measure, one might wonder whether a conditional risk parity portfolio can be regarded as a maximum Sharpe ratio (MSR). From a theoretical perspective, the MSR is appealing because it is mean-variance efficient, while the RP portfolio lies in general below the efficient frontier, unless all assets have the same Sharpe ratios and pairwise correlations (see Maillard et al. [2010]). In practice, it is known that the MSR is plagued by the large uncertainty in expected return estimates, which has a strong negative impact on the out-of-sample performance (see Jorion [1986], Kan and Zhou [2007], and DeMiguel et al. [2009]). Our results show that the MSR portfolio has a noticeably higher turnover than RP portfolios, a property that reflects the well-known fact that MSR weights have a high sensitivity to expected returns, and, as a consequence, are more impacted by estimation errors in these parameters. In other words, a maximum Sharpe ratio portfolio is a less robust alternative to incorporating expected returns in the construction of a well-diversified policy portfolio compared to a conditional risk parity portfolio such as is defined in this article.

APPENDIX

RISK CONTRIBUTIONS FOR DOWNSIDE RISK MEASURES

We provide here the expressions for the contributions of portfolio constituents to total downside risk. Details of the derivation can be obtained from the authors upon request. For a risk measure R_t, the contributions are defined as:

$$c_t^R(\boldsymbol{w}_t) = \boldsymbol{w}_t \odot \frac{\partial R_t}{\partial \boldsymbol{w}_t}$$

where $\frac{\partial R_t}{\partial w_t}$ is the $N \times 1$ gradient vector of R_t.

For Gaussian semi-volatility, we have:

$$c_t^{\mathrm{GSV}}(\boldsymbol{w}_t) = \frac{\mathrm{GSV}_t}{\sigma_{Pt}^2}\boldsymbol{w}_t \odot \Sigma_t \boldsymbol{w}_t + \frac{\ell_{Pt}\mathcal{N}(-\ell_{Pt}) - n(\ell_{Pt})}{\mathrm{GSV}_t}\boldsymbol{w}_t \odot \left[\sigma_{Pt}\mu_t - \ell_{Pt}\Sigma_t \boldsymbol{w}_t\right]$$

where $\mu_{Pt} = \boldsymbol{w}_t' \mu_t$ and $\ell_{Pt} = \mu_{Pt}/\sigma_{Pt}$.

For Gaussian VaR at the confidence level α (equal to 99%), we have:

$$c_t^{\mathrm{GVaR}}(\boldsymbol{w}_t) = \boldsymbol{w}_t \odot \left[-\mu_t + \mathcal{N}^{-1}(\alpha)\tfrac{1}{\sigma_{Pt}}\boldsymbol{\Sigma}_t \boldsymbol{w}_t\right]$$

For non-Gaussian VaR at the level α, the partial derivatives are given by:

$$\frac{\partial \mathrm{NGVaR}_t}{\partial w_t} = \frac{\partial \mathrm{GVaR}_t}{\partial w_t} + \left[\mathrm{NGVaR}_t - \mathrm{GVaR}_t\right] \times \frac{1}{\sigma_{Pt}}\boldsymbol{\Sigma}_t \boldsymbol{w}_t + \sigma_{Pt}\left[\frac{q^2-1}{6}\frac{\partial \mathrm{Sk}_t}{\partial w_t} + \frac{q^3-3q}{24}\frac{\partial \mathrm{Ku}_t}{\partial w_t} - \frac{2q^3-5q}{18}\mathrm{Sk}_t \frac{\partial \mathrm{Sk}_t}{\partial w_t}\right]$$

with:

$$\frac{\partial \mathrm{Sk}_t}{\partial \boldsymbol{w}_t} = \frac{1}{\sigma_{Pt}^3}\frac{\partial m_3}{\partial \boldsymbol{w}_t} - \frac{3m_3}{\sigma_{Pt}^5}\boldsymbol{\Sigma}_t \boldsymbol{w}_t \qquad \frac{\partial \mathrm{Ku}_t}{\partial \boldsymbol{w}_t} = \frac{1}{\sigma_{Pt}^4}\frac{\partial m_4}{\partial \boldsymbol{w}_t} - \frac{4m_4}{\sigma_{Pt}^6}\boldsymbol{\Sigma}_t \boldsymbol{w}_t$$

$$\frac{\partial m_3}{\partial \boldsymbol{w}_t} = 3\boldsymbol{M}_{3t}(\boldsymbol{w}_t \otimes \boldsymbol{w}_t) \qquad \frac{\partial m_4}{\partial \boldsymbol{w}_t} = 4\boldsymbol{M}_{4t}(\boldsymbol{w}_t \otimes \boldsymbol{w}_t \otimes \boldsymbol{w}_t)$$

ENDNOTES

This research has been conducted in the context of the "Asset allocation solutions" research chair at EDHEC-Risk Institute, supported by Lyxor Asset Management. We would like to thank Benjamin Bruder, Nicolas Gaussel, and Thierry Roncalli for very useful comments.

[1]Maillard et al. [2010], Clarke et al. [2013], Spinu [2013] and Griveau-Billion et al. [2013] give examples of such numerical algorithms.

[2]See Coqueret and Milhau [2014] for a survey and a comparison of these techniques.

[3]If a dynamic estimate for the entire covariance matrix is needed, one can use an "M-GARCH" model, which is a multivariate extension allowing for time-dependencies in conditional correlations (see Bollerslev et al. [1988]).

[4]The proof of this result is available from the authors upon request.

[5]With N assets, there are $N(N+1)(N+2)/6$ co-skewnesses and $N(N+1)(N+2)(N+3)/24$ co-kurtosis to estimate, two numbers that quickly grow with N. Martellini and Ziemann [2010] develop an application of shrinkage techniques to the estimation of higher order comoments when N is large. With $N = 2$, however, we do not expect robustness to be a concern.

[6]The proof of this result is available from the authors upon request.

[7]Dividends data, as stock returns, are obtained from CRSP.

[8]With longer prediction horizons, the R^2 and the t-statistics would be higher (see Goetzmann and Jorion [1993] and Valkanov [2003]). However, as argued by Valkanov [2003], the high *t*-statistics obtained at long horizons should not be taken at face value, given that the usual significance levels obtained from the normal distribution are not valid here.

[9]The mean reversion phenomenon might be slow, however, as indicated by the strong autocorrelation of the yield (see Fama and Bliss [1987]).

[10]There is no particular justification for choosing the bond Sharpe ratio as opposed to the stock Sharpe ratio. Note that this choice is irrelevant for the composition of the MSR portfolio computed with prior estimates, but it does have an impact on the weights computed with shrinkage estimators.

[11]The speed of mean reversion is a decreasing function of the autoregression coefficient. In the limit case of a coefficient equal to 1, the speed of mean reversion is zero and the process is a random walk, with no tendency to revert to a particular value.

REFERENCES

Baker, M., R. Greenwood, and J. Wurgler. "The Maturity of Debt Issues and Predictable Variation in Bond Returns." *Journal of Financial Economics*, Vol. 70, No. 2 (2003), pp. 261-291.

Barry, C.B. "Portfolio Analysis under Uncertain Means, Variances, and Covariances." *Journal of Finance*, Vol. 29, No. 2 (1974), pp. 515-522.

Best, M.J., and R.R. Grauer. "On the Sensitivity of Mean-variance-Efficient Portfolios to Changes in Asset Means: Some Analytical and Computational Results." *Review of Financial Studies*, Vol. 4, No. 2 (1991), pp. 315-342.

Bollerslev, T. "Generalized Autoregressive Conditional Heteroskedasticity." *Journal of Econometrics*, Vol. 31, No. 3 (1986), pp. 307-327.

Bollerslev, T., R.F. Engle, and J.M. Wooldridge. "A Capital Asset Pricing Model with Time-varying Covariances." *Journal of Political Economy*, Vol. 96, No. 1 (1988), pp. 116-131.

Boudt, K., B. Peterson, and C. Croux. "Estimation and Decomposition of Downside Risk for Portfolios with Non-Normal Returns." *Journal of Risk*, Vol. 11, No. 2 (2008), pp. 79-103.

Campbell, J., A. Lo, A. MacKinlay, and R. Whitelaw. *The Econometrics of Financial Markets*. Princeton, NJ: Princeton University Press, 1997.

Clarke, R., H. de Silva, and S. Thorley. "Risk Parity, Maximum Diversification, and Minimum Variance: An Analytic Perspective." *The Journal of Portfolio Management*, Vol. 39, No. 3 (2013), pp. 39-53.

Cochrane, J.H., and M. Piazzesi. "Bond Risk Premia." *The American Economic Review*, Vol. 95, No. 1 (2005), pp. 138-160.

Coqueret, G., and V. Milhau. "Estimating Covariance Matrices for Portfolio Optimization." ERI Scientific Beta White Paper, 2014.

Cornish, E.A., and R.A. Fisher. "Moments and Cumulants in the Specification of Distributions." *Revue de l'Institut International de Statistique*, Vol. 5, No. 4 (1938), pp. 307-320.

DeMiguel, V., L. Garlappi, and R. Uppal. "Optimal versus Naive Diversification: How Inefficient Is the 1/*N* Portfolio Strategy?" *Review of Financial Studies*, Vol. 22, No. 5 (2009), pp. 1915-1953.

Dimson, E., P. Marsh, and M. Staunton. *Triumph of the Optimists: 101 Years of Global Investment Returns*. Princeton, NJ: Princeton University Press, 2008.

Duffee, G. "Information in (and Not in) the Term Structure." *Review of Financial Studies*, Vol. 24, No. 9 (2011), pp. 2895-2934.

Fama, E. "Term Premiums in Bond Returns." *Journal of Financial Economics*, Vol. 13, No. 4 (1984), pp. 529-546.

Fama, E., and K. French. "Dividend Yields and Expected Stock Returns." *Journal of Financial Economics*, Vol. 22, No. 1 (1988), pp. 3-25.

Fama, E.F. and R.R. Bliss. "The Information in Long-maturity Forward Rates." *The American Economic Review*, Vol. 77, No. 4 (1987), pp. 680-692.

Goetzmann, W., and P. Jorion. "Testing the Predictive Power of Dividend Yields." *Journal of Finance*, Vol. 48, No. 2 (1993), pp. 663-679.

Griveau-Billion, T., J.-C. Richard, and T. Roncalli. "A Fast Algorithm for Computing High-dimensional Risk Parity Portfolios." Working paper, 2013.

Hallerbach, W. "Decomposing Portfolio Value-at-Risk: A General Analysis." *Journal of Risk*, Vol. 5, No. 2 (2003), pp. 1-18.

Hodrick, R. "Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement." *Review of Financial Studies*, Vol. 5, No. 3 (1992), pp. 357-386.

Ibbotson, R. "Fundamentals of Asset Allocation." Yale SOM EDHEC-Risk Strategic Asset Allocation and Investment Solutions Seminar, 2013.

Ilmanen, A. "Time-Varying Expected Returns in International Bond Markets." *Journal of Finance*, Vol. 50, No. 2 (1995), pp. 481-506.

——. "Forecasting U.S. Bond Returns." *The Journal of Fixed Income*, Vol. 7, No. 1 (1997), pp. 22-37.

Inker, B. "The Hidden Risks of Risk Parity Portfolios." GMO White Paper, 2010.

Jorion, P. "International Portfolio Diversification with Estimation Risk." *Journal of Business*, Vol. 58, No. 3 (1985), pp. 259-278.

——. "Bayes-Stein Estimation for Portfolio Analysis." *Journal of Financial and Quantitative Analysis*, Vol. 21, No. 3 (1986), pp. 279-292.

Joslin, S., M. Priebsch, and K.J. Singleton. "Risk Premiums in Dynamic Term Structure Models with Unspanned Macro Risks." *Journal of Finance*, Vol. 69, No. 3 (2014), pp. 1197-1233.

Kan, R., and G. Zhou. "Optimal Portfolio Choice with Parameter Uncertainty." *Journal of Financial and Quantitative Analysis*, Vol. 42, No. 3 (2007), pp. 621-655.

Litterman, R. "Hot Spots and Hedges." Goldman Sachs Risk Management Series, 1996.

Maillard, S., T. Roncalli, and J. Teïletche. "The Properties of Equally Weighted Risk Contribution Portfolios." *The Journal of Portfolio Management*, Vol. 36, No. 4 (2010), pp. 60-70.

Martellini, L., and V. Ziemann. "Improved Estimates of Higher-Order Comoments and Implications for Portfolio Selection." *Review of Financial Studies*, Vol. 2, No. 4 (2010), pp. 1467-1502.

Menzly, L., T. Santos, and P. Veronesi. "Understanding Predictability." *Journal of Political Economy*, Vol. 112, No. 1 (2004), pp. 1-47.

Merton, R. "On Estimating the Expected Return on the Market: An Exploratory Investigation." *Journal of Financial Economics*, Vol. 8, No. 4 (1980), pp. 323-361.

Newey, W.K., and K.D. West. "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica*, Vol. 55, No. 3 (1987), pp. 703-708.

Qian, E. "Risk Parity Portfolios: Efficient Portfolios through True Diversification." *Panagora*, 2005.

——. "On the Financial Interpretation of Risk Contributions: Risk Budgets Do Add Up." *Journal of Investment Management*, Vol. 4, No. 4 (2006), pp. 41-51.

Roncalli, T. "Introducing Expected Returns in Risk Parity Portfolios: A New Framework for Strategic and Tactical Asset Allocation." Working paper. (2013a). Available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2321309.

——. "Introduction to Risk Parity and Budgeting." Chapman and Hall, CRC Financial Mathematics Series, 2013b.

Schachter, B., and S.R. Thiagarajan. "Risk Parity: Rewards, Risks and Research Opportunities." EDHEC-Risk Institute Publication, 2011.

Spinu, F. "An Algorithm for Computing Risk Parity Weights." Working paper, 2013.

Valkanov, R. "Long-Horizon Regressions: Theoretical Results and Applications." *Journal of Financial Economics*, Vol. 68, No. 2 (2003), pp. 201-232.

Vasicek, O. "A Note on Using Cross-sectional Information in Bayesian Estimation of Security Betas." *Journal of Finance*, Vol. 28, No. 5 (1973), pp. 1233-1239.

Welch, I., and A. Goyal. "A Comprehensive Look at the Empirical Performance of Equity Premium Prediction." *Review of Financial Studies*, Vol. 21, No. 4 (2008), pp. 1455-1508.