Will My Risk Parity Strategy Outperform?

Robert M. Anderson* Stephen W. Bianchi†
Lisa R. Goldberg‡

University of California at Berkeley

July 6, 2012§

Forthcoming in *Financial Analysts Journal*

Abstract

We gauge the return-generating potential of four investment strategies: value weighted, 60/40 fixed mix, unlevered and levered risk parity. We have three main findings. First, even over periods lasting decades, the start and end dates of a backtest can have a material effect on results; second, transaction costs can reverse ranking, especially when leverage is employed; third, a statistically significant return premium does not guarantee outperformance over reasonable investment horizons.

Key words: Risk parity, value weighting, fixed mix, leverage, turnover, trading costs, borrowing costs, market frictions, statistical significance, outperformance, Sharpe ratio

*Department of Economics, 530 Evans Hall #3880, University of California, Berkeley, CA 94720-3880, USA, email: `anderson@econ.berkeley.edu`.

†Department of Economics, 530 Evans Hall #3880, University of California, Berkeley, CA 94720-3880, USA, email: `sbianchi@econ.berkeley.edu`.

‡Department of Statistics, University of California, Berkeley, CA 94720-3880, USA, email: `lrg@stat.berkeley.edu`.

§We are grateful to Rodney Sullivan for stimulating conversations about risk parity, leverage and strategy evaluation, and for support and encouragement. We thank Bernd Scherer and four anonymous referees for careful reviews and insightful comments that improved this article, and we thank Günter Schwarz for educating us about the history of risk parity, and for providing us with early references on the subject. We acknowledge the participants of the Spring 2012 Q Group Seminar, who gave us valuable feedback on our ideas and results. Anderson and Bianchi were supported by the Coleman Fung Chair in Risk Management at UC Berkeley.

1 Past Performance is Not a Guarantee of Future Returns

This familiar disclaimer highlights the fact that a particular investment strategy may work well in some periods and poorly in other periods, limiting the inference that can be drawn from past returns.

The concern is heightened when a proposed investment strategy is backtested using historic data. Consider an investment strategy that can be pursued today using readily available securities. If those securities were not available in the past, then the strategy has no true antecedent. Backtesting must be done using proxies for the securities, and the choice of proxies can have a direct effect on measured returns. In addition, the introduction of new securities can have an indirect effect; a strategy that seems to have been profitable in the past might have been less profitable if the new securities had been available and made the strategy accessible to a broader class of investors. The matter is confounded by the specific attributes of the backtesting period, concerns about statistical significance, and a plethora of metrics used by investors to evaluate strategy performance.

In this article, we consider these issues by carefully examining the historical performance of four simple strategies based on two asset classes: US Equity and US Treasury Bonds.[1] Our study includes a market or value weighted portfolio, which is the optimal risky portfolio in the Capital Asset Pricing Model (CAPM), and a 60/40 mix, which is popular with pension funds and other long horizon investors.

Our study also includes two risk parity strategies. Risk parity attempts to equalize risk contributions across asset classes, Early formulations of risk parity are in Lörtscher (1990) and Kessler and Schwarz (1996).[2] Risk parity has been popular since the 2008 financial crisis, as frustrated investors have struggled to meet return targets by levering low risk or low beta assets, and it is sufficiently mainstream to be featured in the *Wall Street Journal*.[3] A diverse collection of risk parity strategies can be constructed by varying asset classes, grouping schemes, and risk estimates.[4]

An essential element of risk parity is leverage, and it is leverage that distinguishes the two risk parity strategies in our study. An unlevered risk parity strategy tends to have relatively low risk and consequently relatively low expected return, so a risk parity strategy must be levered in order to have even a remote chance of achieving a typical return target.[5] The notion that levering a low-risk portfolio might be worthwhile dates

[1]Our simple, two asset class strategies, which involve no market timing and no security selection, can be used as benchmarks to evaluate more complex strategies that are used in practice.

[2]In Lörtscher (1990) and Kessler and Schwarz (1996), risk parity strategies are known as "equal risk benchmarks."

[3]Dagher (February 6, 2012) discusses the long-term outlook for risk parity strategies.

[4]For example Qian (2005) considers the implications of including asset correlations in risk parity weights. Chaves et al. (2011) consider a broader collection of asset classes, and they also consider risk parity in the context of other low-risk strategies.

[5]There is a large and growing literature on low-risk investing. Sefton et al. (2011) give a broad discus-

back to Black et al. (1972), which provides empirical evidence that the risk-adjusted returns of low beta equities are higher than what is predicted by the CAPM. Black (1972) introduces a zero-beta portfolio, which is considered by some to be the antecedent of risk parity. Nearly four decades later, Frazzini and Pedersen (2010) developed a compelling theory of leverage aversion in which risk parity emerges as a dominant strategy, and this dominance is supported by the empirical study in Asness, Frazzini and Pedersen (2012). However, our results do not support this dominance.

We find that performance depends materially on the backtesting period. For example, in our 85-year Long Sample, 1926–2010, if we assume borrowing was at the risk-free rate[6] and there were no trading costs, the levered risk parity strategy had the highest cumulative return. However the outperformance was not uniform across relatively long subperiods. For example, in our 37-year Post-War Sample, 1946–1982, both the value weighted and 60/40 strategies had higher cumulative returns than the risk parity strategies did.

We find that performance depends materially on assumptions made about market frictions. Since we do not know how the availability of modern financing would have affected markets during the early part of our study period,[7] we extrapolate borrowing costs from recent experience, and we base trading costs on conventional wisdom. We find that market frictions were a substantial drag on performance of the levered risk parity strategy. For example, in our 85-year Long Sample, 1926–2010, after adjusting for transaction costs,[8] both the value weighted and 60/40 strategies had higher cumulative returns than the levered risk parity strategy did. In other words, the ranking based on cumulative return was reversed after adjustment for market frictions. This reversal may be explained by the high degree of leverage in the levered risk parity strategy. The ranking based on cumulative return in Asness, Frazzini and Pedersen (2012) is also reversed. This reversal may be explained by the adjustment for market frictions, and by the fact that the

sion of the topic, and Scherer (2011) attributes the empirically observed outperformance of the market by a particular low-risk (minimum variance) strategy to Fama-French factors. Cowan and Wilderman (2011) provide a rational explanation for the low risk anomaly and Baker et al. (2011) provide a behavioral explanation. Clarke et al. (2011) analyze the connection between minimum variance and low beta strategies.

[6]Throughout this article, the risk-free rate is proxied by the 90-Day T-Bill Rate.

[7]For a liquid asset class such as US Treasury bonds, futures may be the cheapest way to finance the levered position. However, US Treasury futures have been traded in a liquid market only since the 1980s. So it is impossible to conduct a fully empirical study of risk parity that begins early in the twentieth century because we don't know how a futures-financed risk parity strategy would have performed during the Great Depression. We can instead estimate what it would have cost to finance the leverage through more conventional borrowing, but small differences in assumptions about the cost of borrowing have major effects on the estimated returns of a levered risk parity strategy, precisely because the strategy involves such a high degree of leverage. Moreover, because the introduction of liquid US Treasury futures markets presumably reduced the cost of financing a levered risk parity, it may have induced changes in asset returns that would have tended to offset the savings achieved through lower financing costs.

[8]Specifically, borrowing is at the 3-Month Euro-Dollar Deposit Rate starting in 1971, and is equal to the risk-free rate plus sixty basis points before 1971. Turnover-induced trading costs are 1% during the period 1926–1955, .5% during the period 1956–1970 and .1% during the period 1971–2010.

Asness, Frazzini and Pedersen (2012) strategy contains lookahead bias, and is therefore uninvestable.

We find that a statistically significant risk premium may be far from a guarantee of outperformance in practical situations. Under the unrealistic, but nevertheless widely adopted, assumption that the underlying processes possess some strong form of stationarity, the high volatility of security returns poses two closely related practical problems:

- The confidence intervals on the returns of a strategy are very wide, even with many decades of data. Thus, it is rarely possible to demonstrate with conventional statistical significance that one strategy dominates another.

- Even if we were reasonably confident that one strategy achieved higher expected returns than another without incurring extra risk, it would be entirely possible for the weaker strategy to outperform over periods of several decades, certainly beyond the investment horizon of most individuals and even perhaps of institutions like pension funds or endowments.

We find that performance depends on the measure. Over the Long Sample, unlevered risk parity had the highest Sharpe Ratio and the lowest expected return. When unlevered risk parity was levered to have the same volatility as the value weighted portfolio, transaction costs reduced its Sharpe Ratio and its cumulative return was less than the return of the 60/40 and value weighted strategies.[9] Therefore, the empirical observation that levered risk parity outperforms the market in an idealized setting may be explained, at least in part, by the fact that an idealized setting does not include market frictions.

2 Study Outline and Rationale for Some of Our Assumptions

Strategies: We evaluate four strategies based on two asset classes: US Equity and US Treasury Bonds. The strategies are value weighted, 60/40, unlevered risk parity and levered risk parity. Unlevered risk parity is a fully invested strategy weighted so that ex post risk contributions coming from the asset classes are equal. If we lever this strategy to match the ex post volatility of the value weighted portfolio we obtain levered risk parity. Weights in the risk parity strategies depend on volatility estimates, which are based on three-year rolling windows. The strategies are rebalanced monthly. The data and formulas required to replicate our results are in Appendices A and B.

Study Periods: We evaluate the four strategies over an 85-year Long Sample, 1926–2010, and four subperiods. The 20-year Pre-1946 Sample, 1926–1945, which included the

[9]Chaves et al. (2011), comment that realistic borrowing costs might affect the Sharpe ratio: "... it is unclear whether their [unlevered risk parity] Sharpe ratios would remain the same after financing costs."

Great Depression and World War II, was plagued by deflationary shocks and inflationary spikes. Equity markets were uneven during the 37-year Post-War Sample, 1946–1982. This period included spikes in inflation and high interest rates that translated into poor bond performance. The 18-year Bull Market Sample, 1983–2000, included a huge bond rally and the game-changing emergence of the technology industry. The ten-year period that began with the bursting of the DotCom bubble felt turbulent, although it was much calmer than the initial years of the study period.

Transaction Costs: We evaluate the four strategies in each period under three sets of assumptions about transaction costs. The base case assumes borrowing was at the risk-free rate and turnover-induced trading incurred no penalty. The middle case assumes borrowing was at the 3-Month Euro-Dollar Deposit Rate starting in 1971, and was at the risk-free rate plus sixty basis points before 1971. The rationale for this stems from Naranjo (2009), which concludes that investors employing futures borrow at LIBOR rates on average. Since LIBOR rates are available beginning only in 1987, Eurodollar deposit rates are available beginning in 1971, and 3-Month LIBOR and 3-Month Euro-Dollar Deposit Rates track one another closely over the period of overlap, we opted to use Eurodollar deposit rates in our study. The average spread of Eurodollar deposit rates over the risk-free rate during the period 1971–2010 is 100 basis points, so we conservatively assumed a borrowing rate of 60 basis points above the risk-free rate during the 1926–1970 period.

The final case retains borrowing assumptions from the middle case, and adds turnover-induced trading costs of 1% during the period 1926–1955, .5% during the period 1956–1970 and .1% during the period 1971–2010. The details of our turnover estimates and associated penalties are in Appendix C.

Statistical Significance: Confidence in parameters and strategy outperformance is estimated with a non-parametric bootstrap that is described in Appendix D.

Connection to Existing Literature: The data and three of our four strategies: value weighted, 60/40 and unlevered risk parity, are identical to the data and similarly named strategies in the Long Sample in Asness, Frazzini and Pedersen (2012), and our performance estimates match to a high degree of precision. Unlike the levered risk parity strategy in Asness, Frazzini and Pedersen (2012), ours is conditional: it is rebalanced so that its ex post volatility over a three-year window matches the ex post volatility of the value weighted strategy at each rebalancing date. The levered risk parity strategy in Asness, Frazzini and Pedersen (2012) is unconditional: it employs a constant scale factor chosen to match the ex post volatility of the value weighted strategy over the entire study period. Comparing Asness, Frazzini and Pedersen (2012, Figure 1) to our Figure 1, the cumulative return of the unconditional (and uninvestable) levered risk parity strategy was roughly double the cumulative return of the conditional version over the Long Sample.[10]

3 The Specific Start and End Dates of a Backtest Can Have a Material Effect on the Results

Figure 1 shows cumulative returns to the four strategies over the period 1926–2010. Levered risk parity had the highest return by a factor of three. However, the performance was uneven, as shown in Figure 2, where the eight-and-a-half decade study period is broken into four substantial subperiods.

On the basis of cumulative return, levered risk parity prevailed during the the Pre-1946 Sample and the Last 10 Years. Despite its relatively low volatility, even unlevered risk parity beat the value weighted and 60/40 strategies in the most recent period. During the post-war period from 1946 to 1982, both the 60/40 and value weighted strategies outperformed risk parity. Between 1982 and 2000, levered risk parity, 60/40 and value weighted strategies tied for first place.

4 Transaction Costs Can Negate Apparent Outperformance

4.1 Borrowing Costs

In the studies discussed in Section 3, we financed the levered risk parity strategy at the 90-Day T-Bill Rate, but that is not possible in practice. The studies in Naranjo (2009) indicate that in the most recent decade, LIBOR is a more realistic estimate of the implicit interest rate at which investors can lever using futures. Because it is available over a longer period, we use the US 3-Month Euro-Dollar Deposit Rate as a proxy for LIBOR.[11] We repeat the studies in Section 3 replacing the 90-Day T-Bill Rate with the 3-Month Euro-Dollar Deposit Rate rate starting in 1971, and using 90-Day T-Bill Rate plus 60 basis points in the prior period 1926–1970. Because the levered risk parity strategy involves substantial leverage, the effect of this relatively small change in borrowing rate on the return is magnified.

[10] Asness, Frazzini and Pedersen (2012, page 58) find that their unconditional levered risk parity, when financed at LIBOR rates, outperformed 60/40 and value weighted strategies over the Long Sample. They assert that they "obtained similar results by choosing k_t [the factor that scales the strategy to the target volatility level] to match the conditional volatility of the benchmark at the time of portfolio formation." We find that conditional risk parity performs substantially less well than unconditional risk parity, and underperforms 60/40 in the Long Sample when realistic borrowing and trading costs are taken into effect. We also find that unconditional risk parity and 60/40 are virtually tied in the Long Sample when realistic borrowing, or borrowing and trading, costs are taken into account.

[11] Over the period when the US 3-Month Euro-Dollar Deposit Rate and 3-month LIBOR are both available, they track each other very closely, with LIBOR being about 10 basis points higher on average.

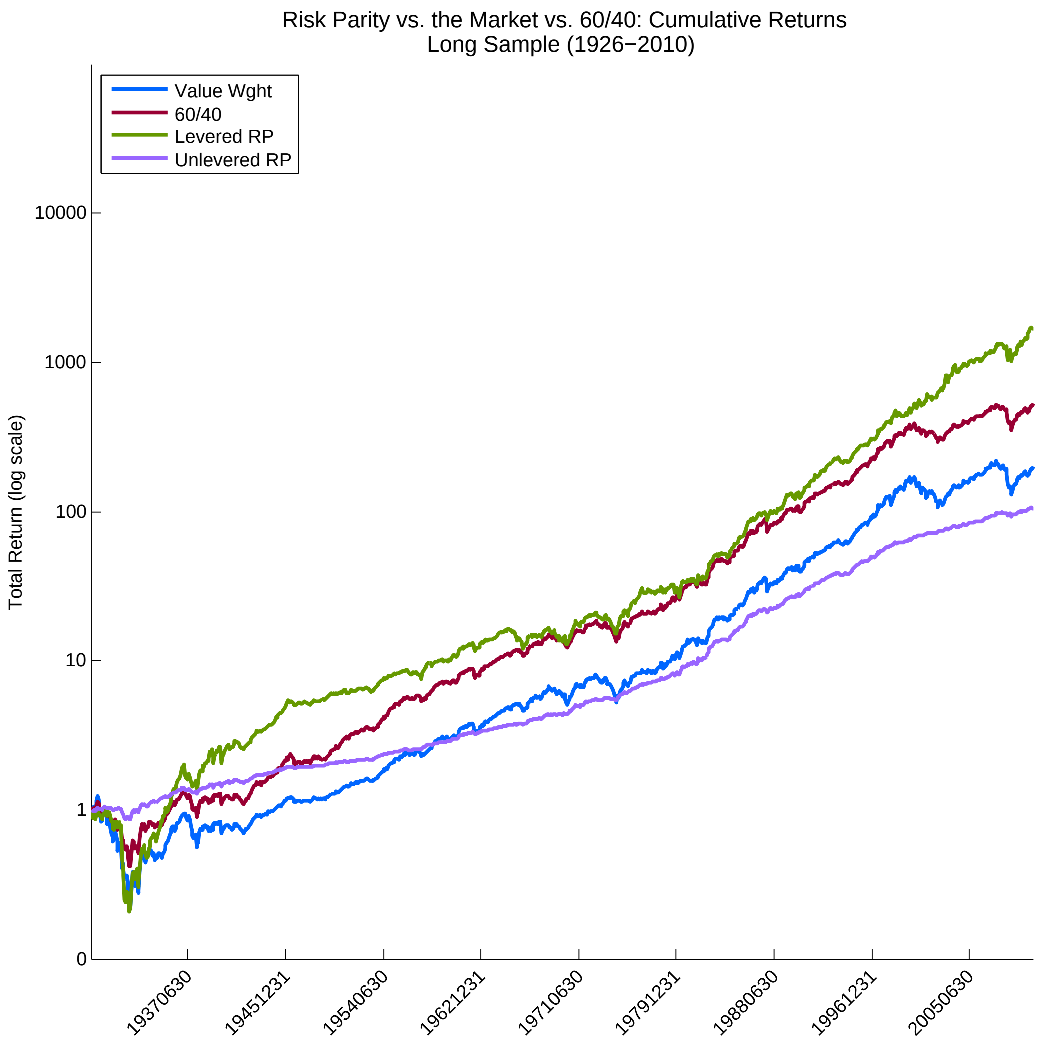

Figure 1: Continuously compounded return to four strategies based on US Equity and US Treasury Bonds over the period 1926–2010. The levered risk parity strategy was financed at the 90-Day T-Bill Rate.

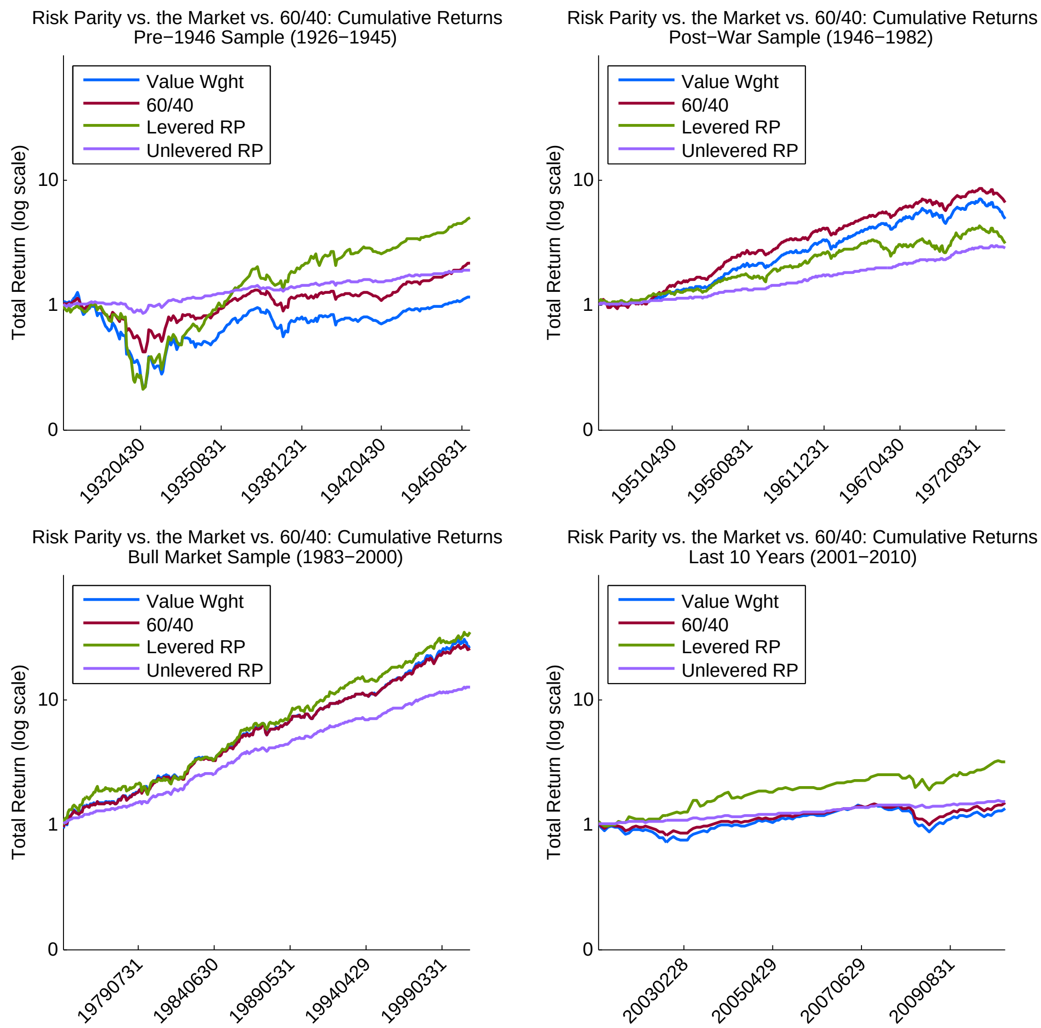

Figure 2: Continuously compounded return to four strategies based on US Equity and US Treasury Bonds over 4 subperiods. The levered risk parity strategy was financed at the at the 90-Day T-Bill Rate. The results depend materially on the evaluation period.

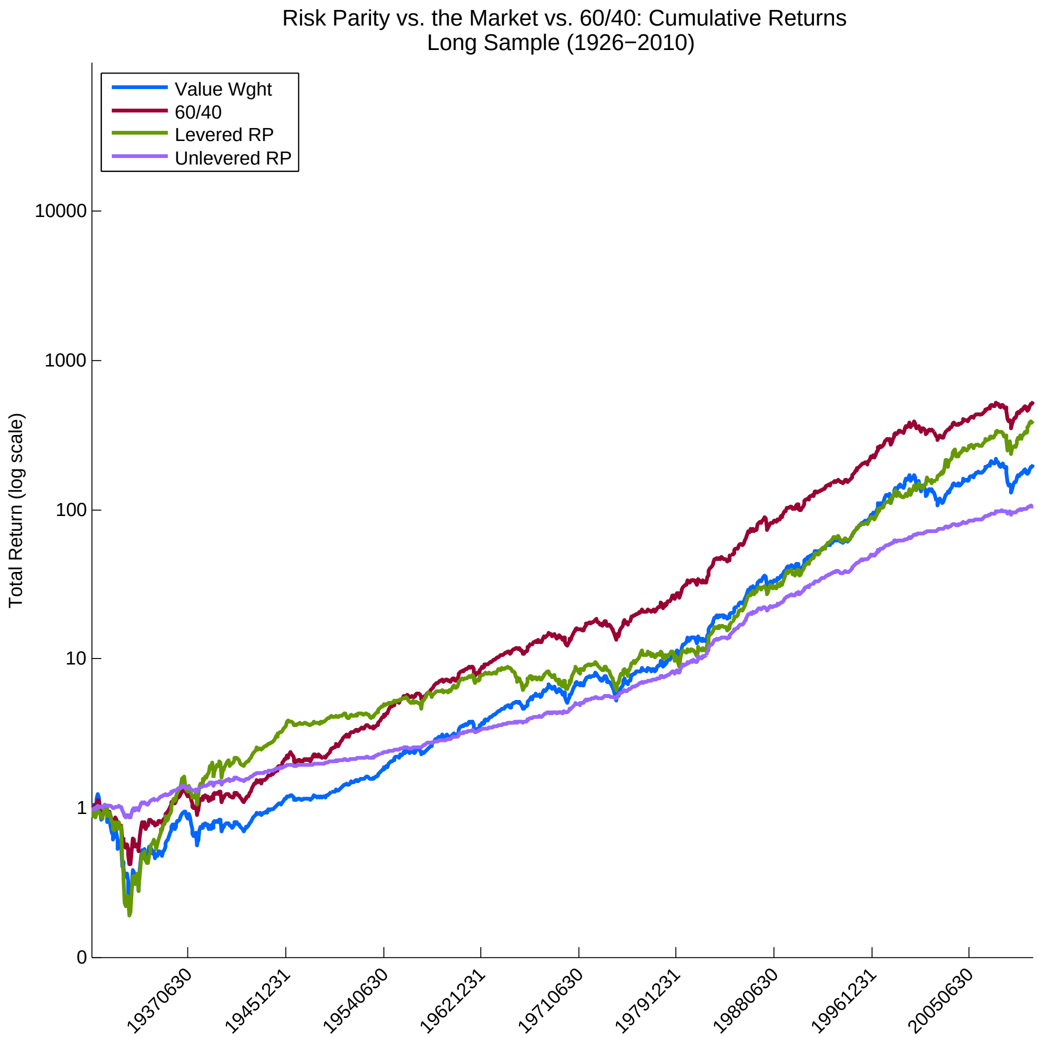

Figure 3: Continuously compounded return to four strategies based on US Equity and US Treasury Bonds over the period 1926–2010. The levered risk parity strategy was financed at the a 3-Month Euro-Dollar Deposit Rate. A comparison with Figure 1 shows the magnitude of the performance drag.

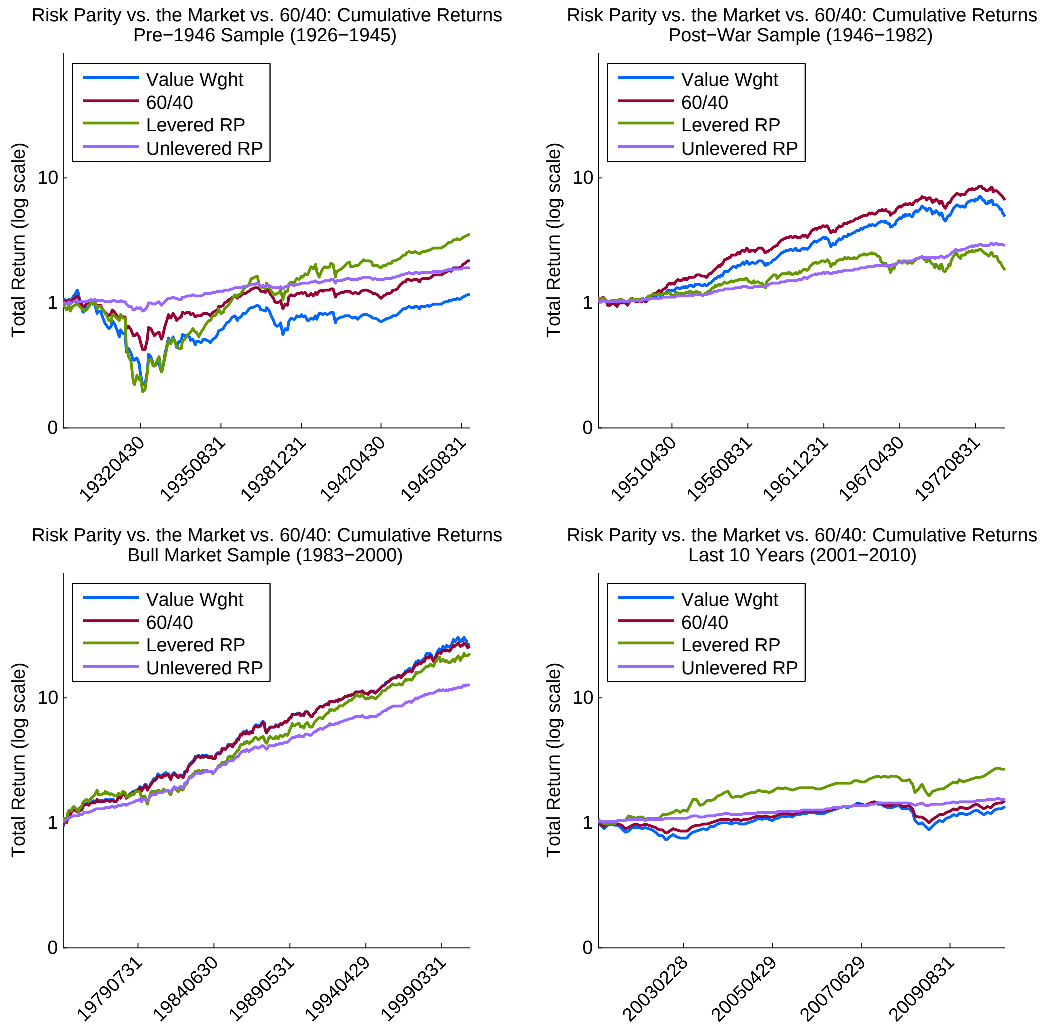

Figure 4: Continuously compounded return to four strategies based on US Equity and US Treasury Bonds over 4 subperiods. The levered risk parity strategy was financed at the 3-Month Euro-Dollar Deposit Rate. A comparison with Figure 2 shows the magnitude of the performance drag, which was most severe in the Post-War sample.

In this experiment, the 60/40 strategy had a slightly higher return than levered risk parity over the long horizon, 1926–2010. This is shown in Figure 3. This reverses the ranking based on cumulative return when borrowing is at the risk-free rate, and it reverses the ranking based on cumulative return in Asness, Frazzini and Pedersen (2012).

The breakdown in Figure 4 is consistent with the assertion that levered risk parity outperforms in turbulent periods and not otherwise. But the data are insufficient to decide on a purely statistical basis whether this assertion has any credence.

4.2 Trading Costs

Value weighted strategies require rebalancing only in response to a limited set of events, for example, new issues and redemptions of bond and shares. The risk parity and 60/40 strategies require additional rebalancing in response to price changes, and hence, they have higher turnover rates. Since we do not have data on new issues or redemptions, and since these should affect the four portfolios in a similar way, we measure the turnover in the risk parity and 60/40 strategies resulting from price changes.[12] As suggested by Figure 5, leverage exacerbates turnover, so the trading costs for the levered risk parity are much higher than they are for the unlevered risk parity and 60/40 strategies. However, the data required to determine the precise relationship between turnover and trading costs are not available. So we estimate.[13]

Figure 6 shows the cumulative return to the four strategies over the long horizon. The levered risk parity strategy is financed at the 3-Month Euro-Dollar Deposit Rate. Turnover-induced trading costs are incorporated in the returns to the 60/40 and risk parity strategies. From the perspective of return, 60/40 is the dominant strategy once again. This time, the value weighted and levered risk parity strategies finish in a tie. Figure 7 shows the breakdown into subperiods.

5 Statistical Significance of Findings Needs to be Assessed

Because the volatility of asset return is substantially larger than its expected value, it is difficult to achieve statistical significance in a comparison of investment strategies, even over periods of decades. Table 1 presents P-values[14] for these comparisons. Disregarding trading costs and assuming borrowing was at the risk-free rate, the (annualized monthly

[12]The details of our turnover estimates are in Appendix C.

[13]We assume trading costs are 1% during the period 1926–1955, .5% during the period 1956–1970 and .1% during the period 1971–2010.

[14]In tests of statistical significance tests, a P-value is the probability of obtaining a test statistic at least as extreme as the one that was actually observed, assuming that the null hypothesis is true. In Table 1, the null hypothesis takes one of two forms: either expected return is zero or regression alpha is zero.

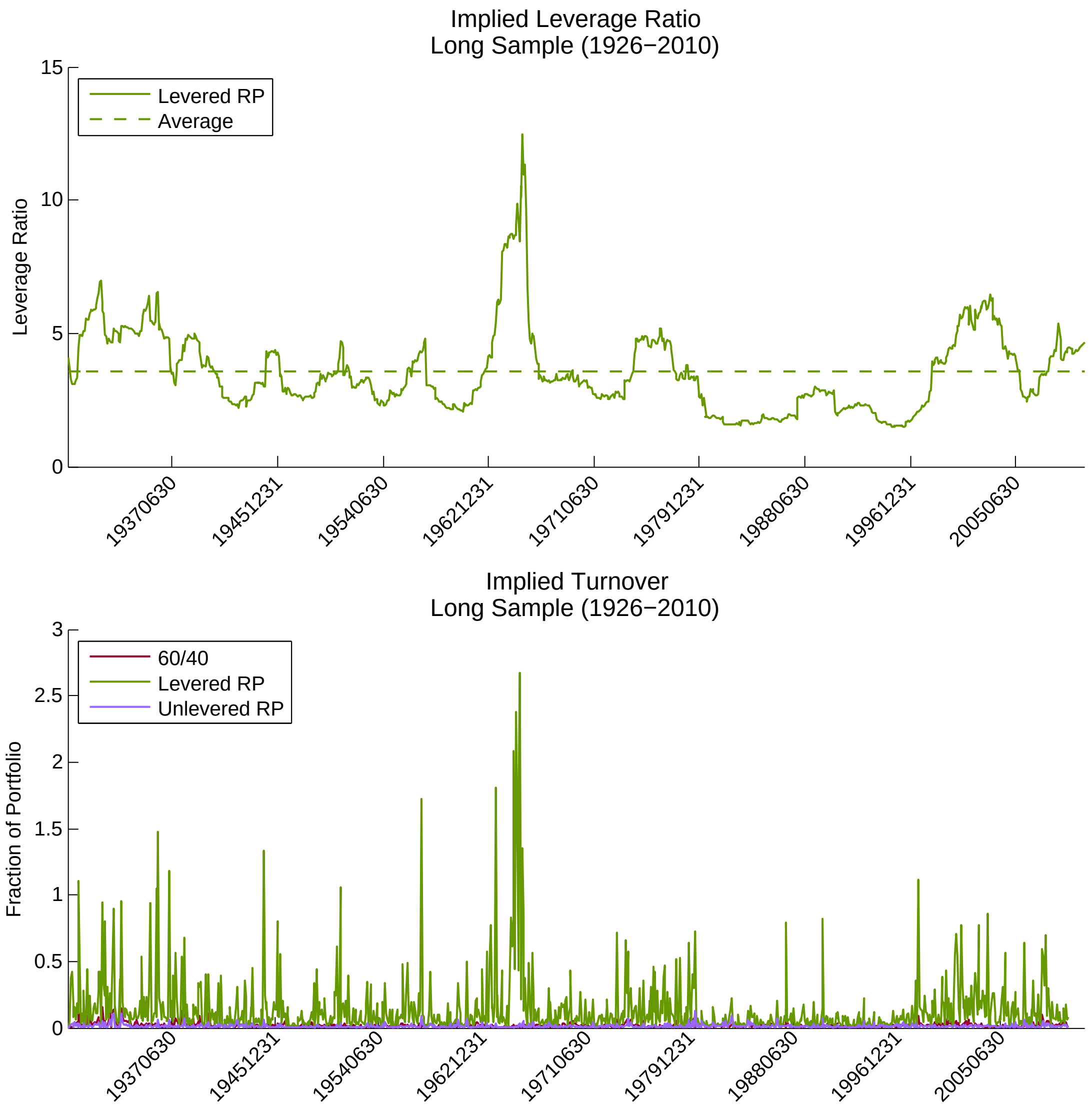

Figure 5: Strategy turnover. The top panel plots the leverage required in order for the estimated volatility of the risk parity strategy to match the estimated volatility of the market at each rebalancing. The average over the entire period was 3.55. The spike in leverage occurred on September 30, 1965, which was a rare moment when bond volatility was relative low (.5%), and both equity volatility (10%) and market weight (72%) were relatively high. The bottom panel shows the turnover of the risk parity and 60/40 strategies at each rebalancing.

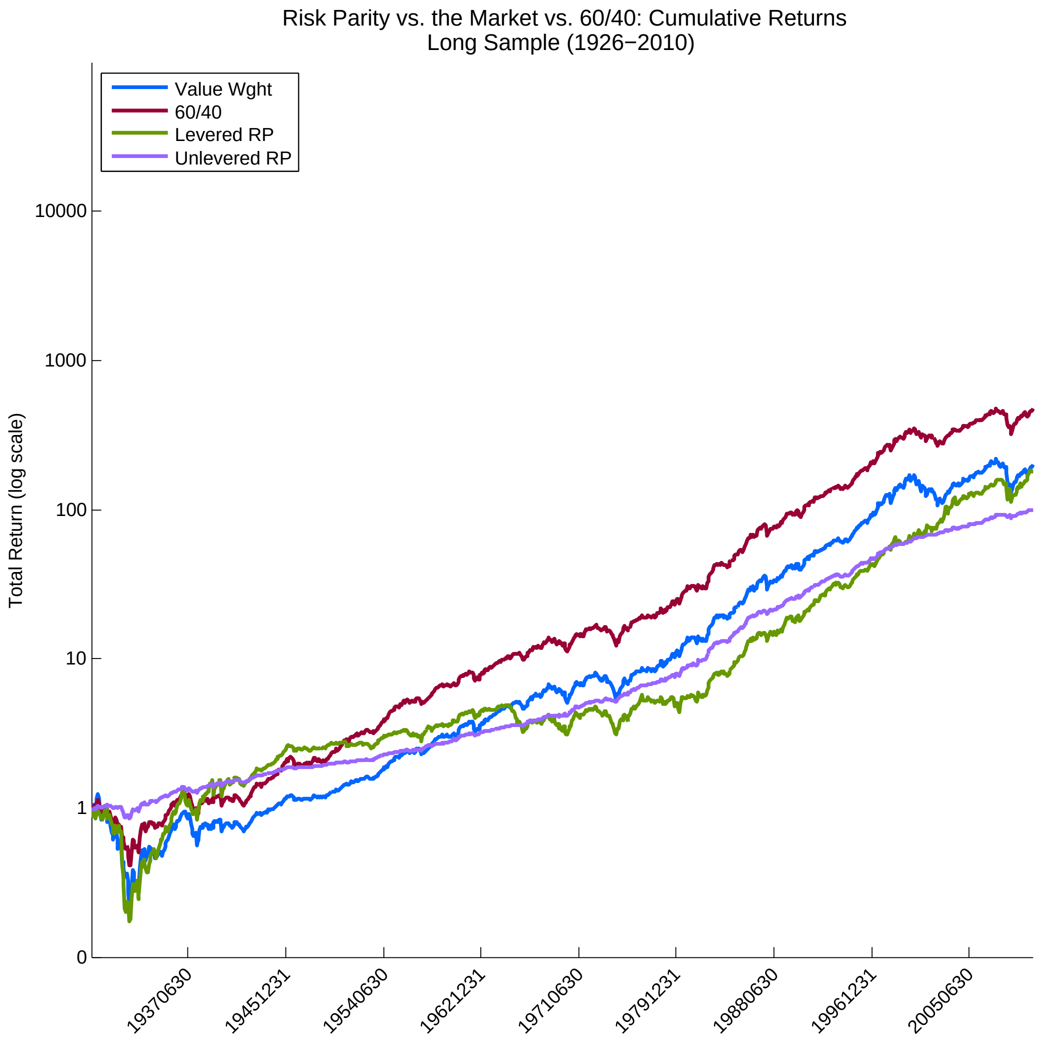

Figure 6: Continuously compounded return to four strategies based on US Equity and US Treasury Bonds over the period 1926–2010. The levered risk parity strategy was financed at the a 3-Month Euro-Dollar Deposit Rate and adjustments are made for turnover. A comparison with Figure 3 shows the magnitude of the performance drag.

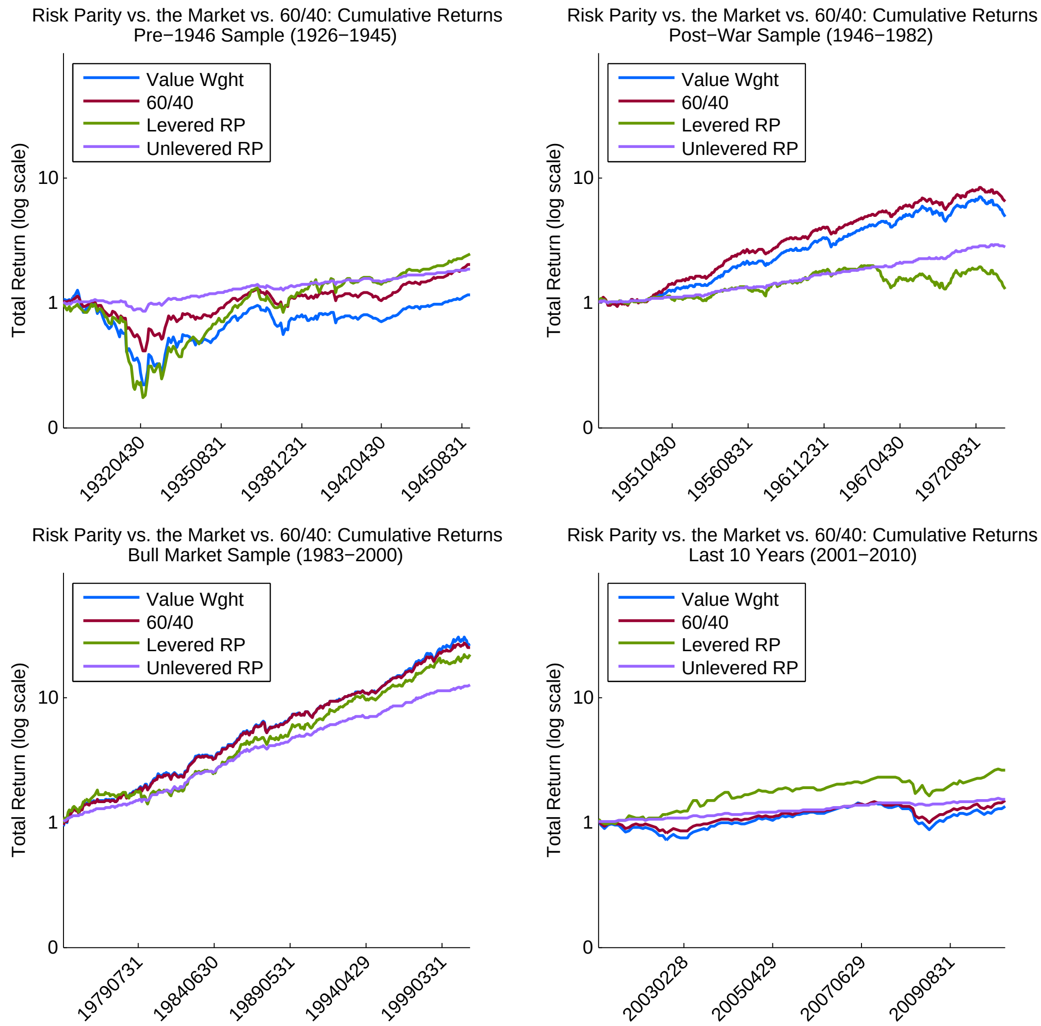

Figure 7: Continuously compounded return to four strategies based on US Equity and US Treasury Bonds over 4 subperiods. The levered risk parity strategy was financed at the 3-Month Euro-Dollar Deposit Rate and adjustments are made for turnover. A comparison with Figure 4 shows the magnitude of the performance drag.

arithmetic) mean return of levered risk parity exceeded that of 60/40 in the 85-year Long Sample by 210 basis points, and the result is statistically significant ($P = 0.03$). However, 60/40 was somewhat less volatile than levered risk parity; taking this into account, the alpha for levered risk parity minus 60/40 just fails to be significant ($P = 0.06$).

Once we take account of borrowing costs that exceed the risk-free rate, the annualized return of levered risk parity exceeded that of 60/40 by only 29 basis points, and is nowhere close to being statistically significant ($P = 0.40$).[15] The alphas were essentially tied. If we also take into account trading costs, 60/40 beat levered risk parity, but the results are not statistically significant. Keep in mind that we are using more than eight decades of data in this analysis, but fail to find statistical significance.

Let's turn the problem around. Suppose we ignore trading costs and assume we can borrow at the risk-free rate. Suppose that, based on our point estimate from our Long Sample, we assume that the expected return of levered risk parity exceeds that of 60/40 by exactly 210 basis points. A bootstrap estimate of the probability that 60/40 will do better than levered risk parity over the next 20 years is 26.8%; over the next 50 years, it is still 17.5%. So even if you ignore borrowing and trading costs, 60/40 has a substantial probability of beating levered risk parity over the next 20 years and the next 50 years.

Of course, even if you do take account of borrowing and trading costs, levered risk parity has a substantial probability of beating 60/40 over the next 20 years and the next 50 years.

6 Risk Profiles

A thorough evaluation of the four investment strategies involves risk as well as return. In this section, we consider the realized Sharpe ratios of the four strategies. Figure 8 shows the strategy Sharpe ratios over 1926–2010, and subperiod Sharpe ratios are in Figure 9. These figures indicate that *unlevered risk parity* has the highest realized Sharpe ratio, with 60/40 coming second.[16] In the Capital Asset Pricing Model (CAPM), the value

[15]There is an apparent conflict between the information in the second panel of Table 1 and the information in Figure 3. Table 1 shows that if we take account of borrowing costs that exceed the risk-free rate but do not adjust for trading costs due to turnover, levered risk parity outperforms 60/40 by a (statistically insignificant) 29 basis points. Figure 3 shows that under the same assumptions, 60/40 outperforms levered risk parity over the 85-year period between 1926–2010. Table 1 reports the arithmetic mean of the monthly returns, which doesn't handle compounding correctly. Figure 3 presents the cumulative returns to the strategies over time, which would correspond to the geometric mean of the monthly returns.

[16]The Sharpe ratios of the levered and unlevered risk parity strategies do not agree, even when borrowing is at the risk-free rate and we ignore trading costs. This is because the leverage is dynamic. The leverage ratio is chosen at each monthly rebalancing so that the conditional ex post volatilities of the levered risk parity and value weighted strategies match. If the levered risk parity strategy were constructed instead with fixed leverage, if borrowing were at the risk-free rate, and if there were no trading costs, the levered risk parity strategy would have the same Sharpe ratio as the unlevered risk parity strategy. Note that fixed leverage is not the same as the unconditional leverage in Asness, Frazzini and Pedersen (2012)

weighted portfolio uniquely maximizes the Sharpe ratio over the feasible set of portfolios with holdings limited to the risky assets. So the results in Figure 8 suggest that the CAPM may not hold.[17]

A consideration that does not depend on the CAPM is the difference between the borrowing rate and the risk-free rate. When that difference is zero, an investor should hold a weighted combination of the risk-free asset and the risky portfolio with the maximum Sharpe ratio. The weights, can be positive, negative or zero. A weighted combination of this type maximizes return for given levels of risk.

However, in the more realistic case when the borrowing rate is higher than the risk-free rate, leverage diminishes the Sharpe ratio. Specifically, for a portfolio with leverage $\lambda > 1$,

$$S_L = S_U - \left(\frac{\lambda - 1}{\lambda}\right)\left(\frac{r_b - r_f}{\sigma}\right) \tag{1}$$

where S_L and S_U are the Sharpe ratios of the otherwise equivalent levered and unlevered portfolios, r_f is the risk-free rate, r_b is the borrowing rate, and σ is the volatilty of the unlevered portfolio. For large leverage,

$$S_L \approx S_U - \frac{r_b - r_f}{\sigma}. \tag{2}$$

When the borrowing rate exceeds the risk-free rate, the efficient frontier is composed of three components, a line segment, an arc of parabola, and a half-line, as depicted schematically in Figure 10. Note that the Sharpe ratio of a levered portfolio on the efficient frontier, which is given by Formula 1, is equal to the slope of the line connecting the portfolio to the risk-free portfolio.

Why did the levered risk parity strategy in Asness, Frazzini and Pedersen (2012) outperform the others after adjusting for financing costs in excess of the risk-free rate, while an analogous adjustment to our levered risk parity strategy caused it to underperform? Asness, Frazzini and Pedersen (2012) match the Long Sample ex post volatility of the levered risk parity to the Long Sample ex post volatility of the value weighted strategy. Of course, this volatility cannot be known in advance, so the levered risk parity strategy in Asness, Frazzini and Pedersen (2012) is not investable.

Table 1 displays standard statistics on the four strategies. The best-performing strategy depends on how an investor weights different risk and performance measures. For example, when one has positive skewness, high kurtosis may be desirable, and this combination occurs for the valued weighted and 60/40 strategies. Levered risk parity exhibits negative skewness and high kurtosis. This bad combination can lead to de-leveraging costs, which could further degrade the performance of levered risk parity, but are beyond the scope of this article. These observations suggest an alternative to leverage aversion as an

[17] Markowitz (2005) discusses a simple paradigm where leverage constraints render the market portfolio inefficient in an idealized setting.

explanation for the performance of the frictionless version of levered risk parity: perhaps there is a premium for taking on severe downside risk.[18]

7 Will My Risk Parity Strategy Outperform?

> When the experiments are done, we still have to decide what to believe.
> - Jonah Lehrer

Strategy evaluation is an important part of the investment process. However, since most strategies do not have true antecedents over long horizons, it is generally not possible to construct fully empirical backtests. Therefore, it is important to evaluate a strategy as broadly as possible—over periods of different length and in different market environments. It is essential to account for market frictions, to keep track of the assumptions underlying extrapolations, to estimate statistical significance, and to interpret results in an economic framework.

In this article, we examined a risk parity strategy of the type considered by pension funds, endowments and other long horizon investors who turn to leverage in an attempt to elevate return in a challenging market. Over the 85-year horizon between 1926 and 2010, the levered risk parity strategy we implemented returned substantially more than unlevered risk parity, a 60/40 fixed mix, and a value weighted portfolio. However, there are important caveats. First, levered risk parity underperformed during a relatively long subperiod: the 37-year Post-War Sample, 1946–1982. Second, transaction costs negated the gains over the full 85-year horizon, 1926–2010. Third, return is but one measure of performance. On the basis of risk-adjusted return, or realized Sharpe ratio, unlevered risk parity dominated the study. Other performance measures might lead to different conclusions.

Compelling economic theories of leverage aversion, such as the one in Frazzini and Pedersen (2010), give credence to the idea that levered risk parity may outperform the market over long horizons. However, there are dissenting voices, such as Sullivan (2010), which are also compelling. The studies in this article suggest that risk parity may be a preferred strategy under certain market conditions, or with respect to certain yardsticks. But any inference from our results must take account of the assumptions we made, and the fact that a study over any horizon, even a long one, is a single draw from a random distribution.

[18] We thank an anonymous referees for the remarks in this paragraph.

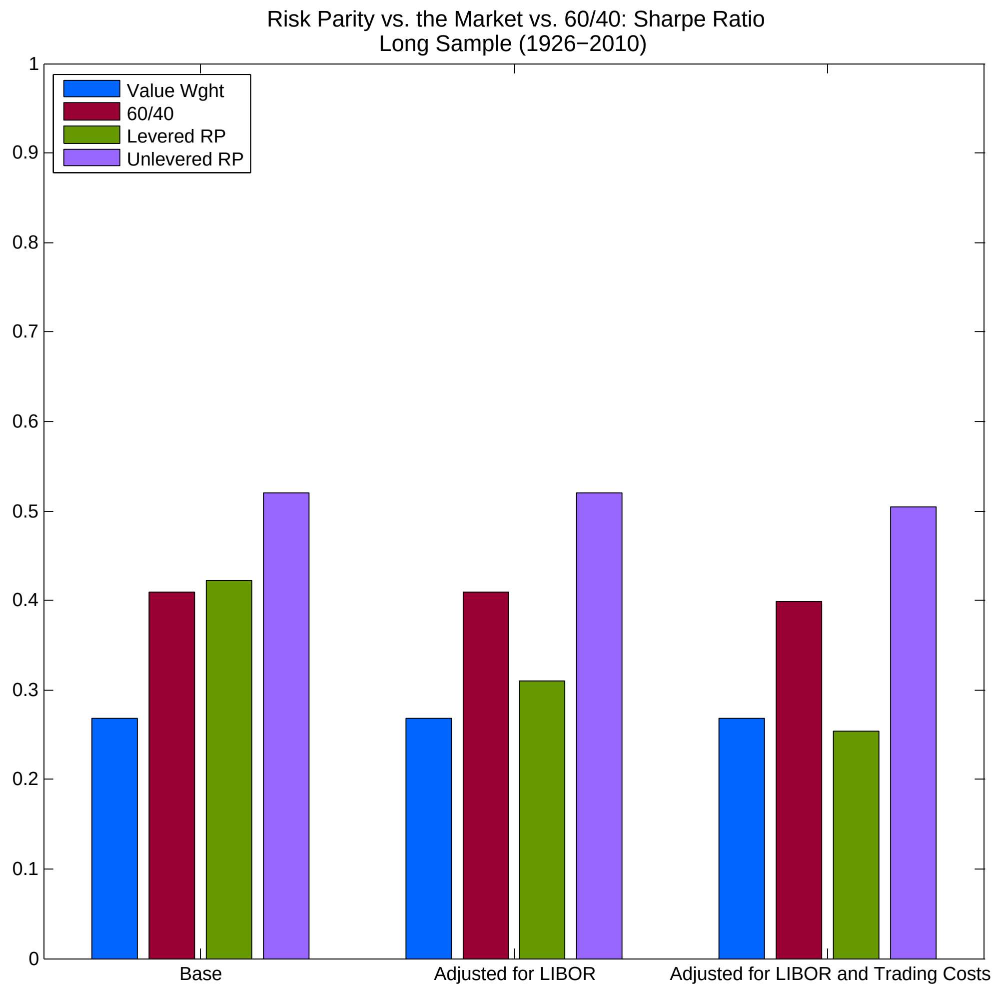

Figure 8: Realized Sharpe ratios for the four strategies over the period 1926–2010. Unlevered risk parity dominates, even before adjustment for market frictions.

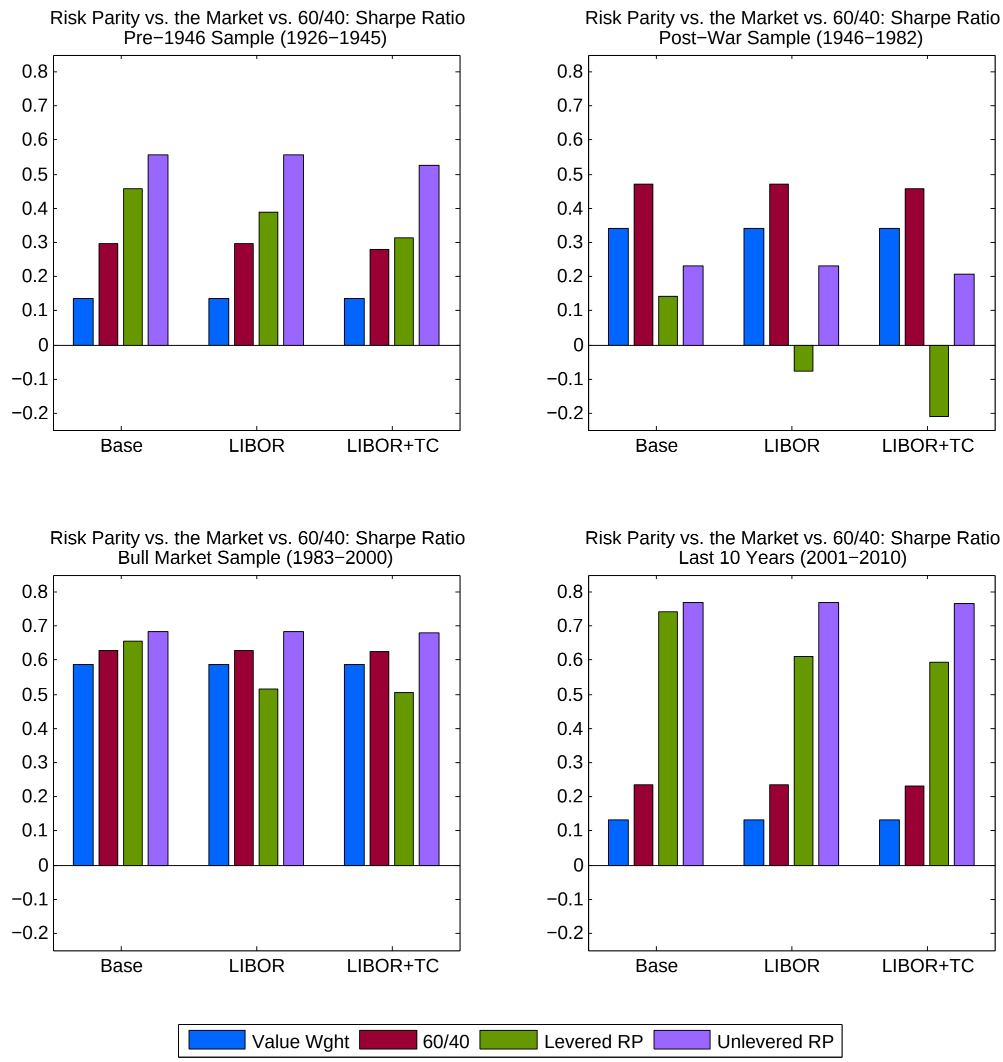

Figure 9: Realized Sharpe ratios for the four strategies over the four subperiods. Apart from the Post-War Sample, Unlevered risk parity dominates, even before adjustment for market frictions.

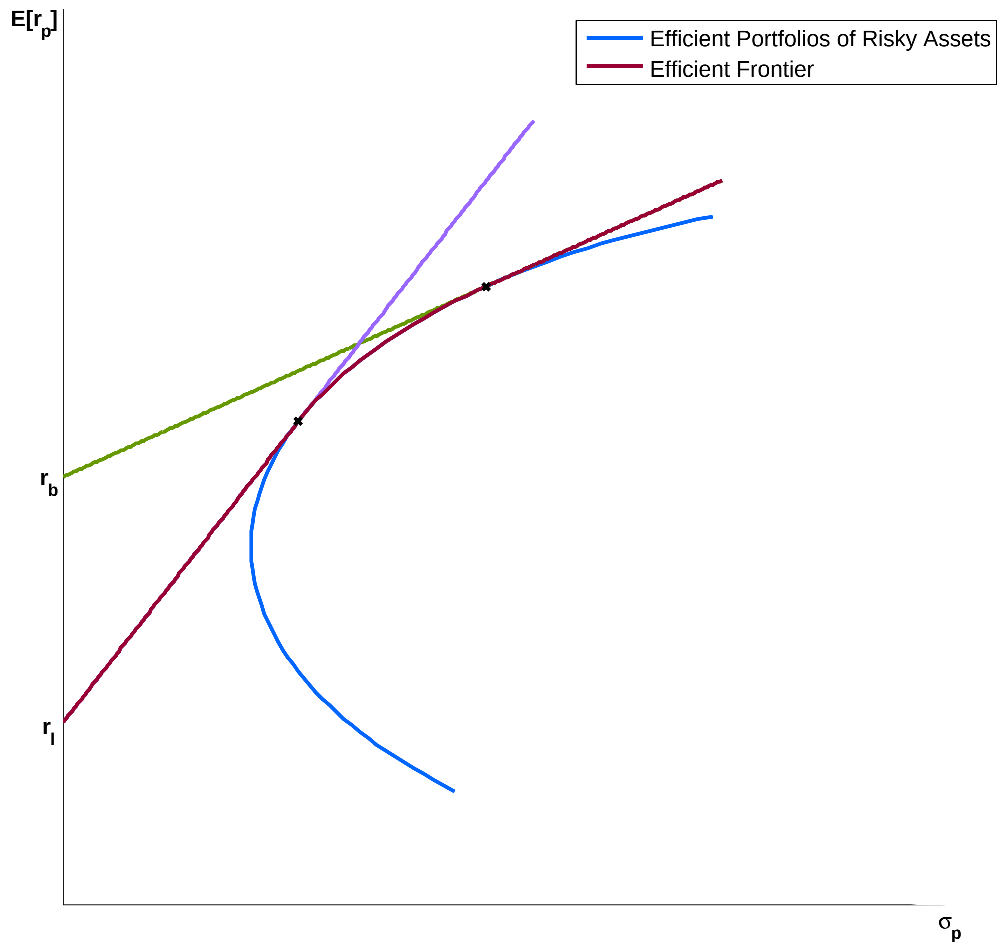

Figure 10: When the rate of borrowing is higher than the risk-free rate, the Capital Market Line in the standard mean-variance diagram has three components. The ex ante Sharpe ratio of a levered portfolio consisting of the market portfolio and cash is lower than the ex ante Sharpe ratio of the market portfolio.

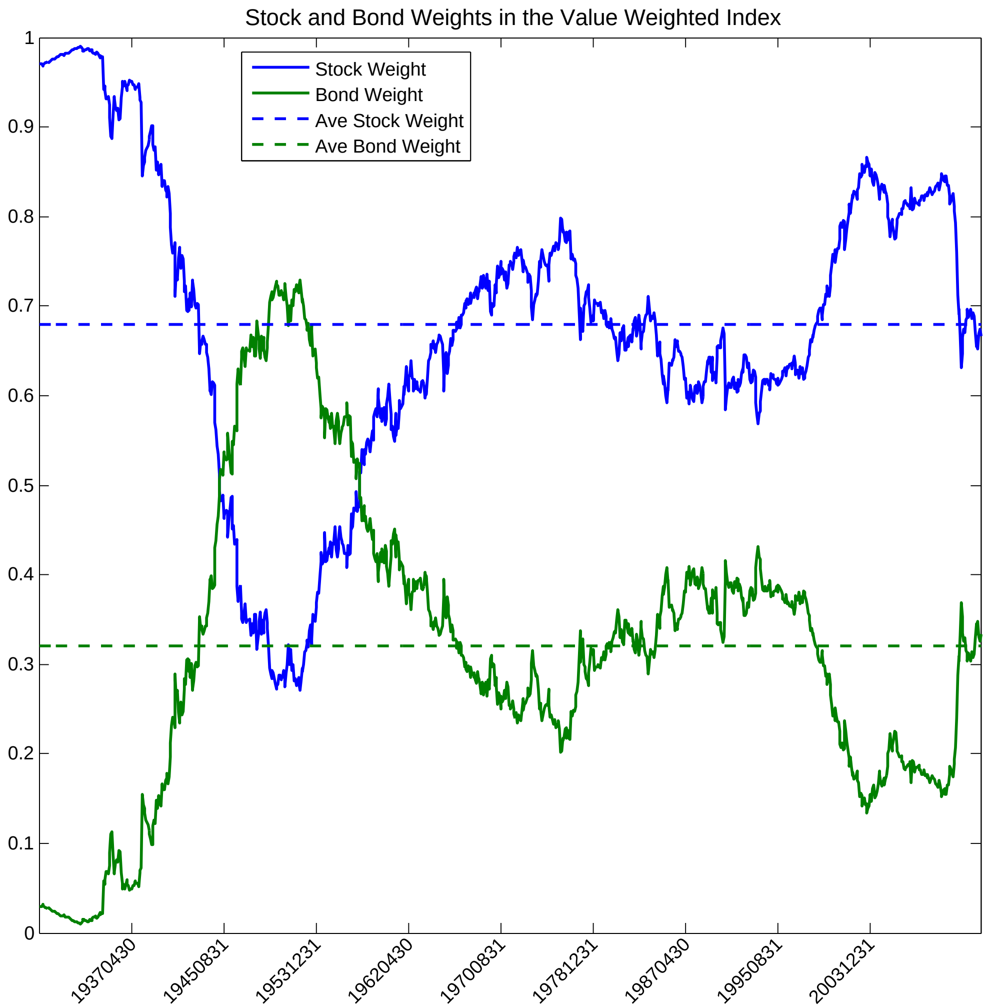

Figure 11: Weights for stocks and bonds implied by market capitalization over the sample period.

Table 1: Risk Parity vs. the Market vs. 60/40 (Historical Performance)

Panel A: Long Sample Stocks and Bonds, 1926-2010 Base Case	Excess Return	*P*-value Excess Return	Alpha	*P*-value Alpha	Volatility	Sharpe Ratio	Skewness	Excess Kurtosis
CRSP Stocks	6.93	0.00			19.05	0.36	0.18	7.44
CRSP Bonds	1.53	0.00			3.28	0.47	0.03	4.74
Value Weighted Portfolio	4.03	0.01			15.04	0.27	0.42	13.58
60/40 Portfolio	4.77	0.00			11.67	0.41	0.20	7.42
Risk Parity (unlevered)	2.21	0.00	1.36	0.00	4.24	0.52	0.07	4.80
Risk Parity (levered)	6.87	0.00	3.53	0.00	16.25	0.42	-0.58	15.54
Risk Parity (levered) minus Val Wght	2.84	0.01	3.53	0.00	10.73	0.26	-0.51	12.42
Risk Parity (levered) minus 60/40	2.10	0.03	1.81	0.06	10.11	0.21	-1.08	13.58

Panel B: Long Sample Stocks and Bonds, 1926-2010 Adjusted for 3M-EDR	Excess Return	*P*-value Excess Return	Alpha	*P*-value Alpha	Volatility	Sharpe Ratio	Skewness	Excess Kurtosis
CRSP Stocks	6.93	0.00			19.05	0.36	0.18	7.44
CRSP Bonds	1.53	0.00			3.28	0.47	0.03	4.74
Value Weighted Portfolio	4.03	0.01			15.04	0.27	0.42	13.58
60/40 Portfolio	4.77	0.00			11.67	0.41	0.20	7.42
Risk Parity (unlevered)	2.21	0.00	1.36	0.00	4.24	0.52	0.07	4.80
Risk Parity (levered)	5.06	0.00	1.70	0.07	16.29	0.31	-0.62	15.47
Risk Parity (levered) minus Val Wght	1.03	0.20	1.70	0.07	10.72	0.10	-0.57	12.50
Risk Parity (levered) minus 60/40	0.29	0.40	-0.02	0.50	10.11	0.03	-1.15	13.68

Panel C: Long Sample Stocks and Bonds, 1926-2010 Adjusted for 3M-EDR and Trading Costs	Excess Return	*P*-value Excess Return	Alpha	*P*-value Alpha	Volatility	Sharpe Ratio	Skewness	Excess Kurtosis
CRSP Stocks	6.93	0.00			19.05	0.36	0.18	7.44
CRSP Bonds	1.53	0.00			3.28	0.47	0.03	4.74
Value Weighted Portfolio	4.03	0.01			15.04	0.27	0.42	13.58
60/40 Portfolio	4.66	0.00			11.67	0.40	0.19	7.39
Risk Parity (unlevered)	2.14	0.00	1.29	0.00	4.24	0.50	0.06	4.80
Risk Parity (levered)	4.15	0.01	0.79	0.24	16.29	0.25	-0.66	15.39
Risk Parity (levered) minus Val Wght	0.11	0.47	0.79	0.24	10.75	0.01	-0.67	13.06
Risk Parity (levered) minus 60/40	-0.51	0.67	-0.81	0.77	10.13	-0.05	-1.22	13.98

Table 1: Performance statistics on the four strategies over the period 1926–2010. In Panel A, the levered risk parity strategy is financed at the 90-Day T-Bill Rate. In Panels B and C, the levered risk parity strategy is financed at the 3-Month Euro-Dollar Deposit Rate. In Panel C, the 60/40 and risk parity strategies are adjusted for turnover.

Appendices

A Data

The results presented in this paper are based on CRSP stock and bond data from January of 1926 through December of 2010. The aggregate stock return is the CRSP value weighted market return (including dividends) from the table *Monthly Stock - Market Indices (NYSE/AMEX/NASDAQ)* – variable name *vwretd*. The aggregate bond return is the face value outstanding (cross-sectionally) weighted average of the unadjusted return for each bond in the *CRSP Monthly Treasury (Master)* table. In this table, the variable name for the unadjusted return is *retnua* and for the face value outstanding is *iout1r*. All bonds in the table are used, provided the values for both *retnua* and *iout1r* are not missing. The value weighted market index is constructed by weighting the aggregate stock return by the total stock market value (variable name *totval*) and the aggregate bond return by the total face value outstanding of all bonds used in the return calculation. Figure 11 plots the stock and bond weights used to estimate the return of the value weighted index.

The proxy for the risk-free rate is the *USA Government 90-day T-Bills Secondary Market* rate, provided by Global Financial Data (http://www.globalfinancialdata.com), covering the period from January of 1926 through December of 2010. The proxy for the cost of financing leverage is the *U.S. 3-Month Euro-Dollar Deposit* rate, downloaded from the Federal Reserve (http://www.federalreserve.gov/releases/h15/data.htm). The 3-Month Euro-Dollar Deposit data is available from January of 1971 through December of 2010. Prior to January of 1971, a constant of 60 basis points is added to the 90-day T-Bill rate.[19]

B Strategies

Rebalancing is monthly.

Value Weighted This is a fully invested strategy that value weights US Equity and US Treasury Bonds.

60/40 This is a fully invested strategy whose capital allocations are 60% US Equity and 40% US Treasury Bonds.

Unlevered Risk Parity This is a fully invested strategy that equalizes ex ante asset class volatilities. The volatility of each asset class is estimated at month end using a 36-

[19]The average difference between the 90-day T-bill Rate and the 3-Month Euro-Dollar Deposit Rate from 1971 through 2010 is roughly 100 basis points. So our estimate of 60 basis points is relatively conservative.

month rolling window of trailing returns. The time-t estimate of volatility for asset class i is given by

$$\hat{\sigma}_{i,t} = \mathrm{std}(r_{i,t-36}, \ldots, r_{i,t-1}).$$

The time-t portfolio weight for asset class i in the unlevered risk parity strategy is given by

$$w_{u,i,t} = \delta_t \hat{\sigma}_{i,t}^{-1},$$

where

$$\delta_t = \frac{1}{\sum_i \hat{\sigma}_{i,t}^{-1}}.$$

Levered Risk Parity This is a levered strategy that equalizes ex ante volatilities across asset classes. The leverage is chosen so that the ex post volatility matches the ex post volatility of the value weighted portfolio at each rebalancing. As in the case of the asset classes, volatility of a strategy is estimated at month end using a 36-month rolling window of trailing returns. The time t estimate of volatility for strategy s is given by

$$\hat{\sigma}_{s,t} = \mathrm{std}(r_{s,t-36}, \ldots, r_{s,t-1}).$$

The leverage ratio required to match the trailing 36-month realized volatility of the value weighted index is the quotient of the volatility estimate for the value weighted portfolio, $\hat{\sigma}_{v,t}$, by the volatility estimate for the unlevered risk parity portfolio, $\hat{\sigma}_{u,t}$:

$$l_t = \frac{\hat{\sigma}_{v,t}}{\hat{\sigma}_{u,t}}.$$

The time-t portfolio weight for asset class i at time t in the levered risk parity strategy is given by

$$w_{l,i,t} = l_t w_{u,i,t}.$$

The return of the levered risk parity portfolio at time t is

$$\begin{aligned} r_{l,t} &= \sum_i w_{u,i,t} r_{i,t} + \sum_i (l_t - 1) w_{u,i,t} (r_{i,t} - r_{b,t}) \\ &= \sum_i w_{u,i,t} r_{i,t} + \sum_i (w_{l,i,t} - w_{u,i,t})(r_{i,t} - r_{b,t}), \end{aligned}$$

where $r_{b,t}$ is the borrowing rate at time t.

Asness, Frazzini and Pedersen (2012) implement an unconditional levered risk parity strategy. The asset class weights in this strategy depend on a time-independent scale factor k chosen so that volatility of excess returns estimated over the entire sample, 1926–2010, matches the volatility of excess returns of the value weighted strategy. To be precise,

$$w_{l.unc,i,t} = k \hat{\sigma}_{i,t}^{-1},$$

$$r^e_{l.unc,t} = r_{l.unc} - r_{f,t},$$

and

$$\sigma = \mathrm{std}(r^e_{l.unc,37}, \ldots, r^e_{l.unc,T}), \tag{3}$$

where σ is a desired target volatility, (which Asness, Frazzini and Pedersen (2012) set to be the realized volatility of the value-weighted portfolio). Here, T is the *last* month in the sample period (i.e. if the sample period is January 1926 through December 2010, then $T = 1020$). Note that the target σ is not known until the end of the period. Moreover, even if σ were set to some constant that were known in 1926, k cannot be computed until the full history through 2010 is known. If k and σ were set to some constants in 1926, then Equation 3 would not be satisfied. Thus, this version of the unconditional levered risk parity is not investable.

The conditional and unconditional levered risk parity strategies differ in other important ways. Consider, for example, their responses to an upward spike in equity volatility. All else equal, both strategies will increase the ratio of capital in bonds to capital in equity, but the conditional strategy will increase its leverage, while the unconditional strategy will decrease its leverage.

C Trading Costs

To estimate trading costs due to turnover, we need to express the change in portfolio weights due to price movements (or returns) over a single period. For any strategy, the time-t return-modified weight to asset i is given by

$$\tilde{w}_{i,t} = \frac{(1 + r_{i,t})w_{i,t-1}}{\sum_j (1 + r_{j,t})w_{j,t-1}}$$

and the turnover required to rebalance the strategy is given by

$$x_t = \sum_j |\tilde{w}_{j,t} - w_{j,t}|.$$

In view of the large and variable leverage implicit in our levered risk parity strategy, we explicitly show the impact of leverage on turnover.

$$x_t = \sum_j |\tilde{w}_{u,j,t}\ell_{t-1} - w_{u,j,t}\ell_t|.$$

Trading costs at time t are then given by

$$c_t = x_t z_t,$$

where (by assumption) z_t is equal to 1% for 1926-1955, 0.5% for 1956-1970, and 0.1% for 1971-2010, and trading cost adjusted returns are given by

$$r'_{l,t} = r_{l,t} - c_t.$$

Table 2: Effect of Alternate Borrowing Cost Assumptions for the Pre-1971 Period

Pre-1971 Borrowing Cost Spread Over T-bills (bp)	Excess Return	*P*-value Excess Return	Alpha	*P*-value Alpha	Volatility	Sharpe Ratio	Skewness	Excess Kurtosis
A. Long sample, adjusted for 3M-EDR (RP - value-weighted)								
25	1.54	0.10	2.22	0.03	10.72	0.14	-0.54	12.41
50	1.17	0.16	1.85	0.05	10.72	0.11	-0.56	12.48
60	1.03	0.20	1.70	0.07	10.72	0.10	-0.57	12.50
75	0.81	0.25	1.48	0.10	10.73	0.08	-0.58	12.54
100	0.44	0.36	1.11	0.17	10.73	0.04	-0.61	12.59
125	0.07	0.47	0.74	0.26	10.74	0.01	-0.63	12.64
B. Long sample, adjusted for 3M-EDR (RP - 60/40)								
25	0.80	0.23	0.50	0.33	10.10	0.08	-1.11	13.53
50	0.43	0.35	0.13	0.45	10.11	0.04	-1.14	13.64
60	0.29	0.40	-0.02	0.50	10.11	0.03	-1.15	13.68
75	0.06	0.47	-0.24	0.58	10.11	0.01	-1.16	13.74
100	-0.30	0.60	-0.61	0.70	10.12	-0.03	-1.19	13.84
125	-0.67	0.73	-0.97	0.81	10.13	-0.07	-1.22	13.93
C. Long sample, adjusted for 3M-EDR and Trading Costs (RP - value-weighted)								
25	0.63	0.29	1.31	0.13	10.74	0.06	-0.64	12.99
50	0.26	0.41	0.94	0.21	10.74	0.02	-0.66	13.04
60	0.11	0.47	0.79	0.24	10.75	0.01	-0.67	13.06
75	-0.11	0.53	0.57	0.31	10.75	-0.01	-0.68	13.09
100	-0.48	0.66	0.20	0.43	10.76	-0.04	-0.70	13.13
125	-0.84	0.76	-0.16	0.56	10.77	-0.08	-0.73	13.17
D. Long sample, adjusted for 3M-EDR and Trading Costs (RP - 60/40)								
25	0.01	0.50	-0.29	0.60	10.12	0.00	-1.18	13.84
50	-0.36	0.64	-0.66	0.73	10.13	-0.04	-1.21	13.94
60	-0.51	0.67	-0.81	0.77	10.13	-0.05	-1.22	13.98
75	-0.73	0.74	-1.03	0.82	10.14	-0.07	-1.24	14.03
100	-1.10	0.84	-1.40	0.89	10.15	-0.11	-1.26	14.11
125	-1.47	0.91	-1.77	0.94	10.16	-0.14	-1.29	14.18

Table 2: Impact of the borrowing cost extrapolation on the risk premium of levered risk parity over the value weighted strategy. 1926–2010. In the top panel, the levered risk parity strategy is financed the 3-Month Euro-Dollar Deposit Rate. In the bottom panel, an additional adjustment is made for trading due to turnover.

D Bootstrap Estimates

In order to reflect the empirical properties of our data, we use a bootstrap to estimate the P-values in Table 1. For a given strategy and evaluation period, suppose we have a sample of T monthly observations of excess return. The excess return reported in Table 1 is the annualized mean. To estimate the P-value for the excess return, we draw 10,000 bootstrap samples of T observations (with replacement) from the empirical distribution. We calculate the mean of each bootstrap sample. The P-value is given by:

$$P = \frac{\#\text{means} <= 0}{N}.$$

The bootstrap procedure for the alpha P-value is different. Suppose

$$R_s = \alpha + \beta R_b + \epsilon,$$

where R_b is the vector of excess returns of a benchmark portfolio (i.e. $R_b = (R_{b,1}, \ldots, R_{b,T})'$), which in our case is the value weighted portfolio, and R_s is the vector of excess returns of a strategy portfolio. A time series regression to estimate alpha and beta generates the residuals:

$$e_t = R_{s,t} - \hat{\alpha} - \hat{\beta} R_{b,t},$$

for $t = 1, \ldots, T$. Next, we draw 10,000 samples (with replacement) of T observations from the empirical distribution of residuals and, for each sample, regenerate the strategy returns as:

$$R_s^* = \hat{\alpha} + \hat{\beta} R_b + \epsilon^*,$$

where ϵ^* is the vector of resampled residuals. Then for each sample, we run a time series regression based on the equation above to get new estimates of alpha ($\hat{\alpha}^*$) and beta ($\hat{\beta}^*$). The P-value for alpha is given by:

$$P = \frac{\#\hat{\alpha}^* <= 0}{N}.$$

The probability estimates in Section 5 are also based on a bootstrap. For example, to calculate the probability that 60/40 will outperform levered risk parity over a 20 year horizon, we draw $10,000$ samples of 240 contemporaneous monthly observations from empirical distribution of the total returns to the 60/40 and levered risk parity portfolios. For each sample, we calculate the cumulative return to each strategy over the 20 year horizon and record the difference $cr_d = cr_{rp} - cr_{60/40}$. The probability estimate is given by:

$$\mathbb{P} = \frac{\#cr_d < 0}{N}.$$

E Two Robustness Checks

We consider the impact of two of the assumptions that are made in this study.

E.1 Rebalancing Horizon

The monthly rebalancing horizon used in our studies is shorter than the horizon typically used by pensions, endowments, and other long-term investors. Figure 12 shows the cumulative return to the four strategies when they are rebalanced annually. The left panel show the base case where borrowing is at the 90-Day T-Bill Rate and no adjustment is made for trading due to turnover. In the right panel, borrowing is at the 3-Month Euro-Dollar Deposit Rate and adjustments are made for trading due to turnover. The results are less dramatic than they are when rebalancing monthly horizon, but they are qualitatively similar.

E.2 Borrowing Cost Assumptions

The US 3-Month Euro-Dollar Deposit Rates used, starting in 1971, as the estimate of the implicit interest rate when levering through futures. In the prior period 1926–1970 when the 3-Month Euro-Dollar Deposit Rate was not available, we extrapolate borrowing costs to be the 90-Day T-Bill Rate plus 60 basis points. This is a conservative estimate since the average spread between the3-Month Euro-Dollar Deposit Rate and the 90-Day T-Bill Rate during the period 1971–2010 is roughly 100 basis points. Table 2 shows the impact of varying the extrapolated spread on strategy performance. We consider spreads ranging between 25 and 125 basis points. After adjustment for turnover-induced trading, the risk premium of levered risk parity over the market had low statistical significance even when a level of 25 basis points is taken to be the spread between the 3-Month Euro-Dollar Deposit Rate and the 90-Day T-Bill Rate between 1926 and 1970.

References

Clifford Asness, Andrea Frazzini and Lasse Heje Pedersen. Leverage Aversion and Risk Parity. *Financial Analysts Journal*, 8(1):47–59, 2012.

Malcolm Baker, Brendon Bradley and Jeffry Wurgler. Benchmarks as Limits to Arbitrage: Understanding the Low Volatility Anomaly. *Financial Analysts Journal*, 78(1):40–54, 2011.

Fischer Black. Capital market equilibrium with restricted borrowing. *Journal of Business*, 45(3):444–455, 1972.

Fischer Black, Michael C. Jensen and Myron Scholes. The Capital Asset Pricing Model: Some Empirical Tests. In Michael C. Jensen, editor, *Studies in the Theory of Capital Markets*, pages 79–121. Praeger Publishers Inc., 1972.

Denis Chaves, Jason Hsu, Feifei Li and Omid Shakernia. Risk Parity Portfolio vs. Other Asset Allocation Heuristic Portfolios. *Journal of Investing*, 20(1):108–118, 2011.

Roger Clarke, Harindra De Silva and Stephen Thorley. Minimum Variance Portfolio Composition. *Journal of Portfolio Management*, 37(2):31–45, 2011.

David Cowan and Sam Wilderman. Rethinking Risk: What the Beta Puzzle Tells us About Investing, 2011. GMO White Paper.

Veronica Dagher. New Allocation Funds Redefine Idea of 'Balance'. *Wall Street Journal*, February 6, 2012.

Andrea Frazzini and Lasse Heje Pedersen. Betting Against Beta, 2010. Working Paper, AQR Capital Management, New York University and NBER Working Paper 16601.

Christoph Kessler and Günter Schwarz. Investment management for private investors at SBC - The benchmark, 1996. Swiss Bank Corporation / The Global 2nd Quarter.

Rudolf Lörtscher. Investment Policy at SBC - basic premises and model, 1990. Swiss Bank Corporation / The Global 10/90.

Harry M. Markowitz. Market Efficiency: A Theoretical Distinction and So What? *Financial Analysts Journal*, September/October:17–30, 2005.

Lorenzo Naranjo. Implied Interest Rates in a Market with Frictions, 2009. Working paper, New York University.

Edward Qian. Risk Parity Portfolios: Efficient Portfolios through True Diversification, 2005. Panagora Research Pape.

Bernd Scherer. A Note on the Returns From Minimum Variance Investing. *Journal of Empirical Finance*, 18:652–660, 2011.

James Sefton, David Jessop, Giuliano De Rossi, Claire Jones and Heran Zhang. Low-Risk Investing, 2011. UBS Investment Research.

Rodney Sullivan. Speculative Leverage: A False Cure for Pension Woes. *Financial Analysts Journal*, 66(3):6–8, 2010.

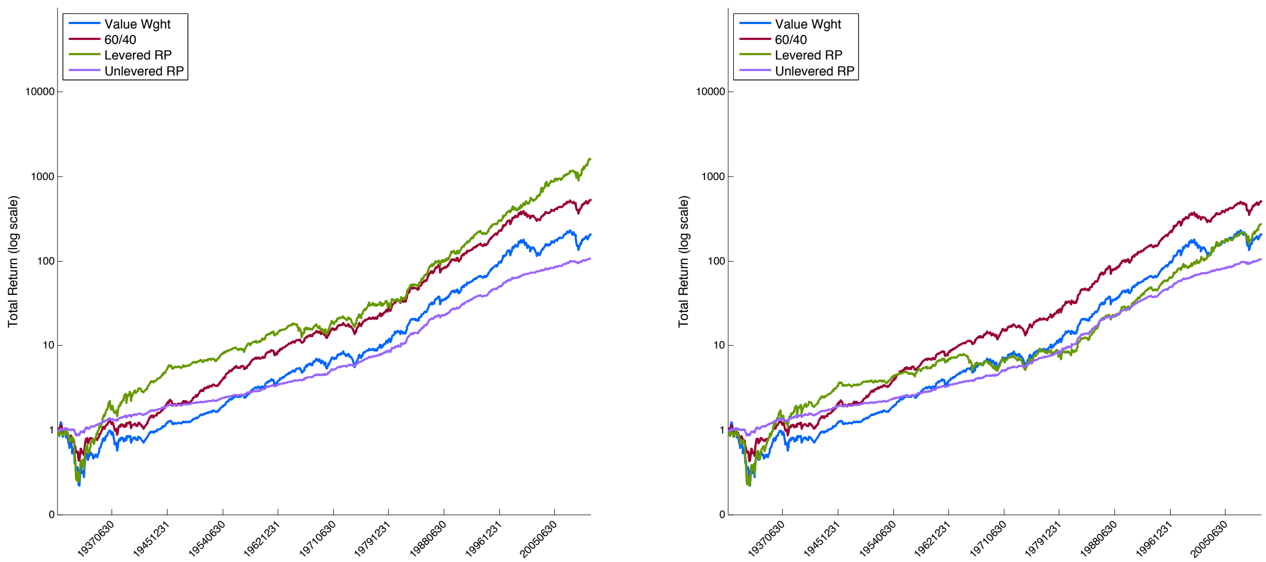

Figure 12: Continuously compounded return to four strategies based on US Equity and US Treasury Bonds over the period 1926–2010 with annual rebalancing. In the left panel, the levered risk parity strategy was financed at the 90-Day T-Bill Rate and no adjustments are made for turnover. In the right panel, the levered risk parity strategy was financed at the 3-Month Euro-Dollar Deposit Rate and adjustments are made for turnover.