Financial Analysts Journal
Volume 68 • Number 6
©2012 CFA Institute

Will My Risk Parity Strategy Outperform?

Robert M. Anderson, Stephen W. Bianchi, CFA, and Lisa R. Goldberg

The authors gauged the return-generating potential of four investment strategies: value weighted, 60/40 fixed mix, and unlevered and levered risk parity. They report three main findings: (1) Even over periods lasting decades, the start and end dates of a backtest can have a material effect on results; (2) transaction costs can reverse ranking, especially if leverage is used; and (3) a statistically significant return premium does not guarantee outperformance over reasonable investment horizons.

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 others, limiting the inference that can be drawn from past returns.

The concern is heightened when a proposed investment strategy is backtested by using historical data. Consider an investment strategy that can be pursued today with readily available securities. If those securities were unavailable in the past, then the strategy has no true antecedent. Backtesting must be conducted with 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 was seemingly profitable in the past might have been less profitable if the new securities had been available and thus 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 our study, we considered these issues by carefully examining the historical performance of four simple strategies based on two asset classes: U.S. equity and U.S. Treasury bonds.[1] We included 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. We also included two risk parity strategies. Risk parity attempts to equalize risk contributions across asset classes; early formulations of risk parity can be found 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, which distinguishes the two risk parity strategies in our study. Because an unlevered risk parity strategy tends to have relatively low risk and thus relatively low expected return, 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 back to Black, Jensen, and Scholes (1972), who provided empirical evidence that the risk-adjusted returns of low-beta equities are higher than the CAPM would predict. Black (1972) introduced a zero-beta portfolio, 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.

■ *Discussion of findings.* We found that performance depends materially on the backtesting period. For example, in our long sample, covering 85 years (1926–2010), if we assumed borrowing at the risk-free rate[6] and 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

Robert M. Anderson is professor of economics, Stephen W. Bianchi, CFA, is a PhD candidate in economics, and Lisa R. Goldberg is professor of statistics at the University of California, Berkeley.

value-weighted and the 60/40 strategies had higher cumulative returns than the risk parity strategies.

We also found that performance depends materially on assumptions about market frictions. Because we could not know how the availability of modern financing would have affected markets during the early part of our study period,[7] we extrapolated borrowing costs from recent experience and based trading costs on conventional wisdom. We found that market frictions were a substantial drag on the performance of the levered risk parity strategy. For example, in our long sample (1926–2010), after adjusting for transaction costs,[8] both the value-weighted and the 60/40 strategies had higher cumulative returns than the levered risk parity strategy. In other words, the ranking based on cumulative return was reversed after the 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) was also reversed. This reversal may be explained by both the adjustment for market frictions and the fact that their strategy contains look-ahead bias and is thus uninvestable.

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

- Because the confidence intervals on the returns of a strategy are very wide, even with many decades of data, 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 such institutions as pension funds or endowments.

Finally, we found 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.

Study Outline and Rationale for Our Assumptions

Before we report our results, we discuss the strategies we used, our sample periods, transaction costs, statistical significance, and our study's connection to the literature.

Strategies. We evaluated four strategies—value weighted, 60/40, unlevered risk parity, and levered risk parity—based on two asset classes: U.S. equity and U.S. Treasury bonds. Unlevered risk parity is a fully invested strategy weighted so that *ex post* risk contributions 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. The weights in our risk parity strategies depended on volatility estimates, which we based on three-year rolling windows. We rebalanced our strategies monthly. The data and formulas required to replicate our results are in Appendix A.

Sample Periods. We evaluated 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 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), a period of inflationary spikes 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 10-year period that began with the bursting of the dot-com bubble seemed turbulent, although it was much calmer than the initial years of our study period.

Transaction Costs. We evaluated the four strategies in each period under three sets of assumptions about transaction costs. The base case assumed borrowing at the risk-free rate and turnover-induced trading with no penalty. The middle case assumed borrowing at the three-month Eurodollar deposit rate starting in 1971 and at the risk-free rate plus 60 bps before 1971. The rationale for these assumptions stems from Naranjo (2009), who concluded that investors using futures borrow at LIBOR, on average. Because LIBOR is available only from 1987 on, Eurodollar deposit rates are available from 1971 on, and three-month LIBOR and the three-month Eurodollar deposit rate track one another closely over the period of overlap, we opted to use Eurodollar deposit rates in our study.

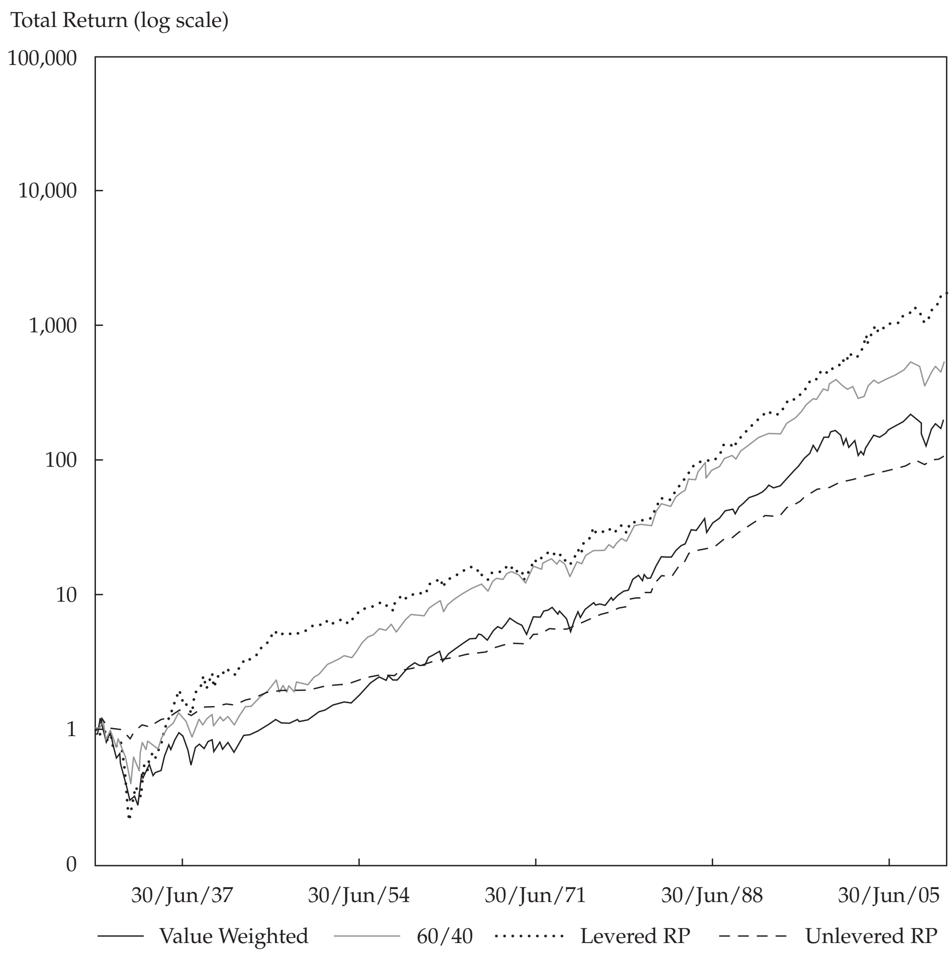
The average spread of Eurodollar deposit rates over the risk-free rate for 1971–2010 is 100 bps, and so we conservatively assumed a borrowing rate of 60 bps above the risk-free rate over 1926–1970.

The final case retained the borrowing assumptions from the middle case and added turnover-induced trading costs of 1% over 1926–1955, 0.5% over 1956–1970, and 0.1% over 1971–2010. The details of our turnover estimates and associated penalties are in Appendix A.

Statistical Significance. We estimated the statistical significance of parameters and strategy outperformance with a nonparametric bootstrap that is described in Appendix A.

Connection to the 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 theirs to a high degree of precision. Unlike their levered risk parity strategy, however, 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. Their levered risk parity strategy is unconditional: It uses a constant scale factor chosen to match the *ex post* volatility of the value-weighted strategy over the entire study period. Comparing Figure 1 in Asness, Frazzini, and Pedersen (2012) with our **Figure 1**, we can see that the cumulative return of our conditional levered risk parity strategy over the long sample was roughly half the cumulative return of their unconditional (and uninvestable) version.[10]

Figure 1. Cumulative Returns for Risk Parity vs. the Market vs. 60/40, 1929–2010



Notes: This figure shows monthly compounded returns to four strategies based on U.S. equity and U.S. Treasury bonds. The levered risk parity strategy was financed at the 90-day T-bill rate. RP indicates risk parity strategies.

Effect of a Backtest's Start and End Dates on Its Results

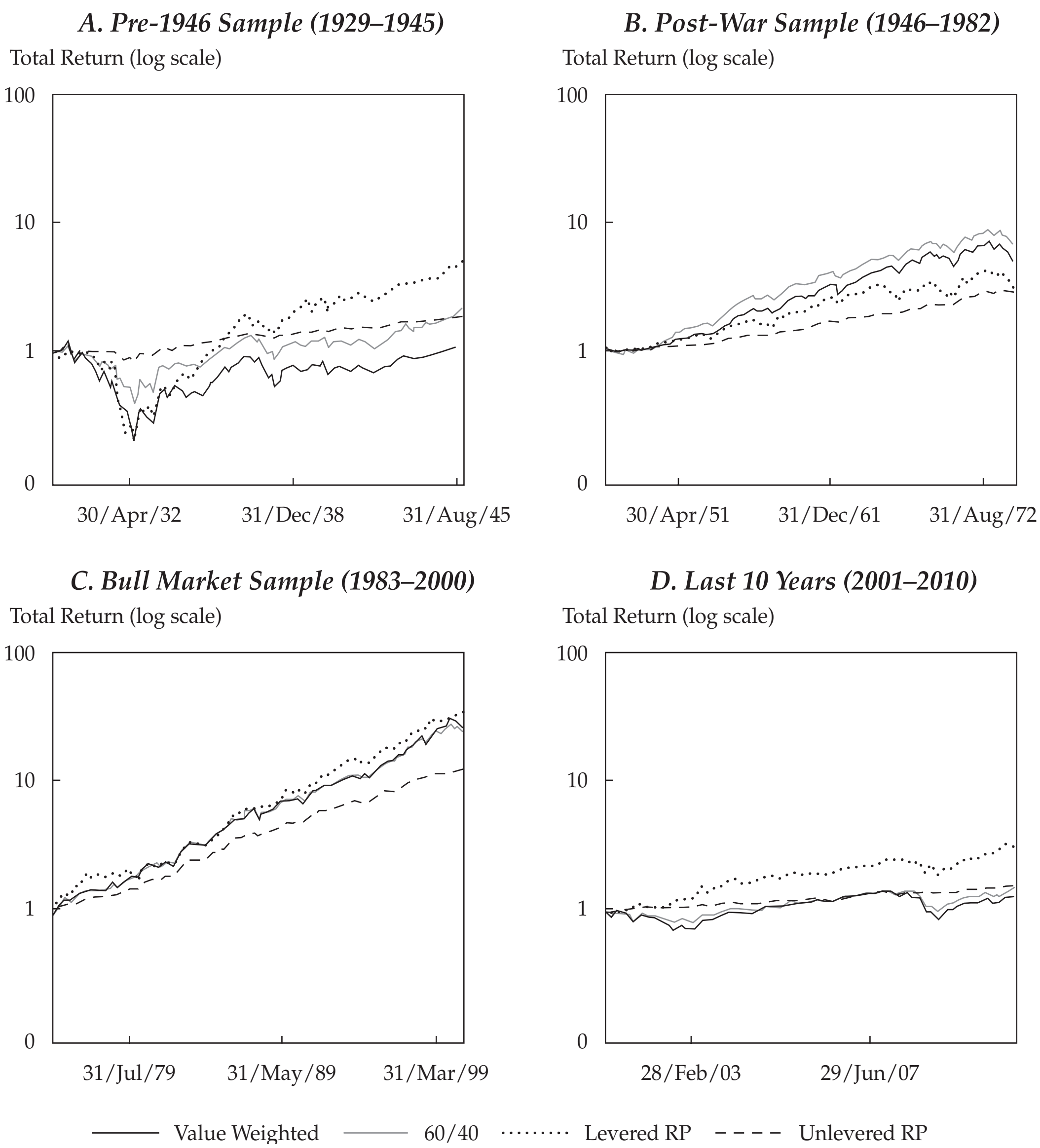
Figure 1 shows cumulative returns for the four strategies over 1926–2010. Levered risk parity had the highest return by a factor of 3. However, the performance was uneven, as shown in **Figure 2**, in which the 85-year sample period is broken into four substantial subperiods.

On the basis of cumulative return, levered risk parity prevailed for both 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 (1946–1982), both the 60/40 and the value-weighted strategies outperformed risk parity. Between 1982 and 2000, the levered risk parity, 60/40, and value-weighted strategies tied for first place.

Effect of Transaction Costs on Apparent Outperformance

Borrowing and trading costs can negate apparent outperformance.

Figure 2. Cumulative Returns for Risk Parity vs. the Market vs. 60/40 over Various Sample Periods



Notes: This figure shows monthly compounded returns to four strategies based on U.S. equity and U.S. Treasury bonds over four subperiods. The levered risk parity strategy was financed at the 90-day T-bill rate. The results depend materially on the evaluation period.

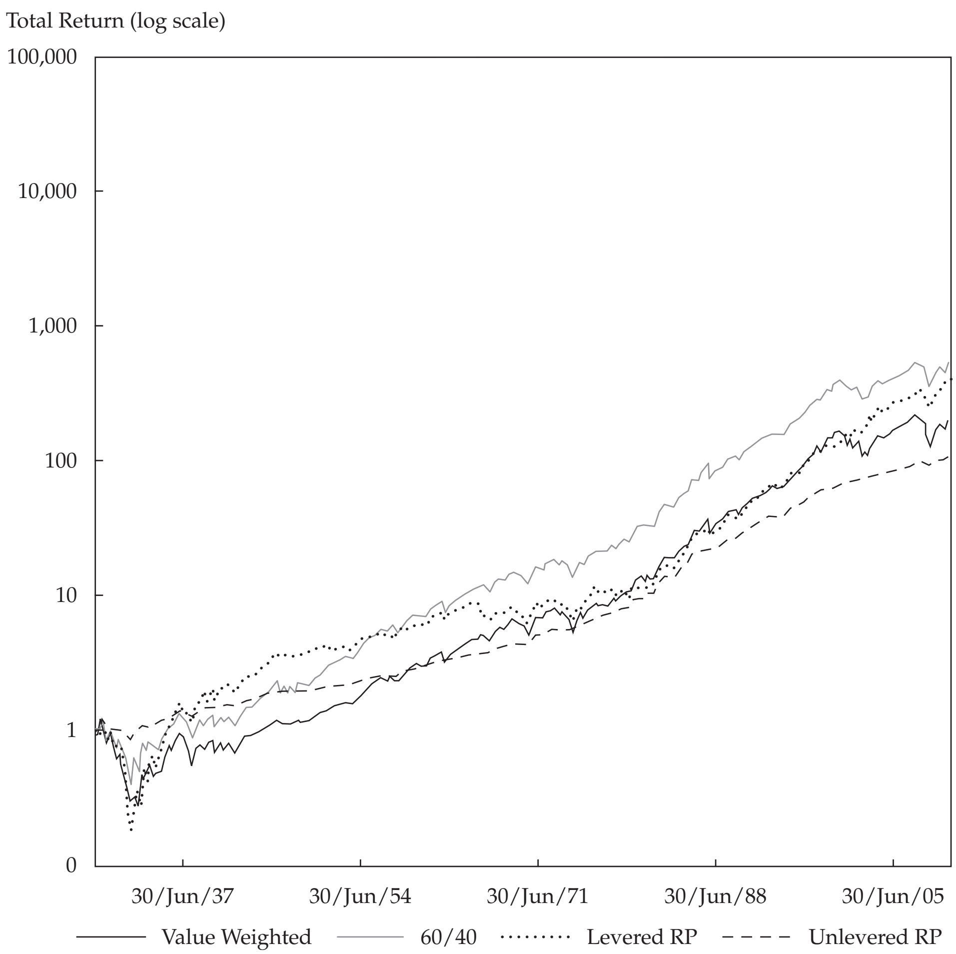
Borrowing Costs. In the experiments discussed in the previous section, we financed the levered risk parity strategy at the 90-day T-bill rate, but that approach is not possible in practice. Naranjo (2009) demonstrated that in the most recent decade, LIBOR has been a more realistic estimate of the implicit interest rate at which investors can lever by using futures. Because it is available over a longer period, we used the U.S. three-month Eurodollar deposit rate as a proxy for LIBOR.[11] We repeated these experiments, replacing the 90-day T-bill rate with the three-month Eurodollar deposit rate starting in 1971 and using the 90-day T-bill rate plus 60 bps for 1926–1970. Because the levered risk parity strategy involves substantial leverage, the effect of this relatively small change in the borrowing rate on the return is magnified.

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

Trading Costs. Value-weighted strategies require rebalancing only in response to a limited set of events—for example, new issues and redemptions of bonds and shares. The risk parity and 60/40 strategies require additional rebalancing in response to price changes and thus have higher turnover rates. Because we had no data on new issues or redemptions, which should affect the four portfolios in a similar way, we measured the turnover in the risk parity and 60/40 strategies resulting from price

Figure 3. Cumulative Returns for Risk Parity vs. the Market vs. 60/40, 1929–2010



Notes: This figure shows monthly compounded returns to four strategies based on U.S. equity and U.S. Treasury bonds. The levered risk parity strategy was financed at the three-month Eurodollar deposit rate. A comparison with Figure 1 shows the magnitude of the performance drag.

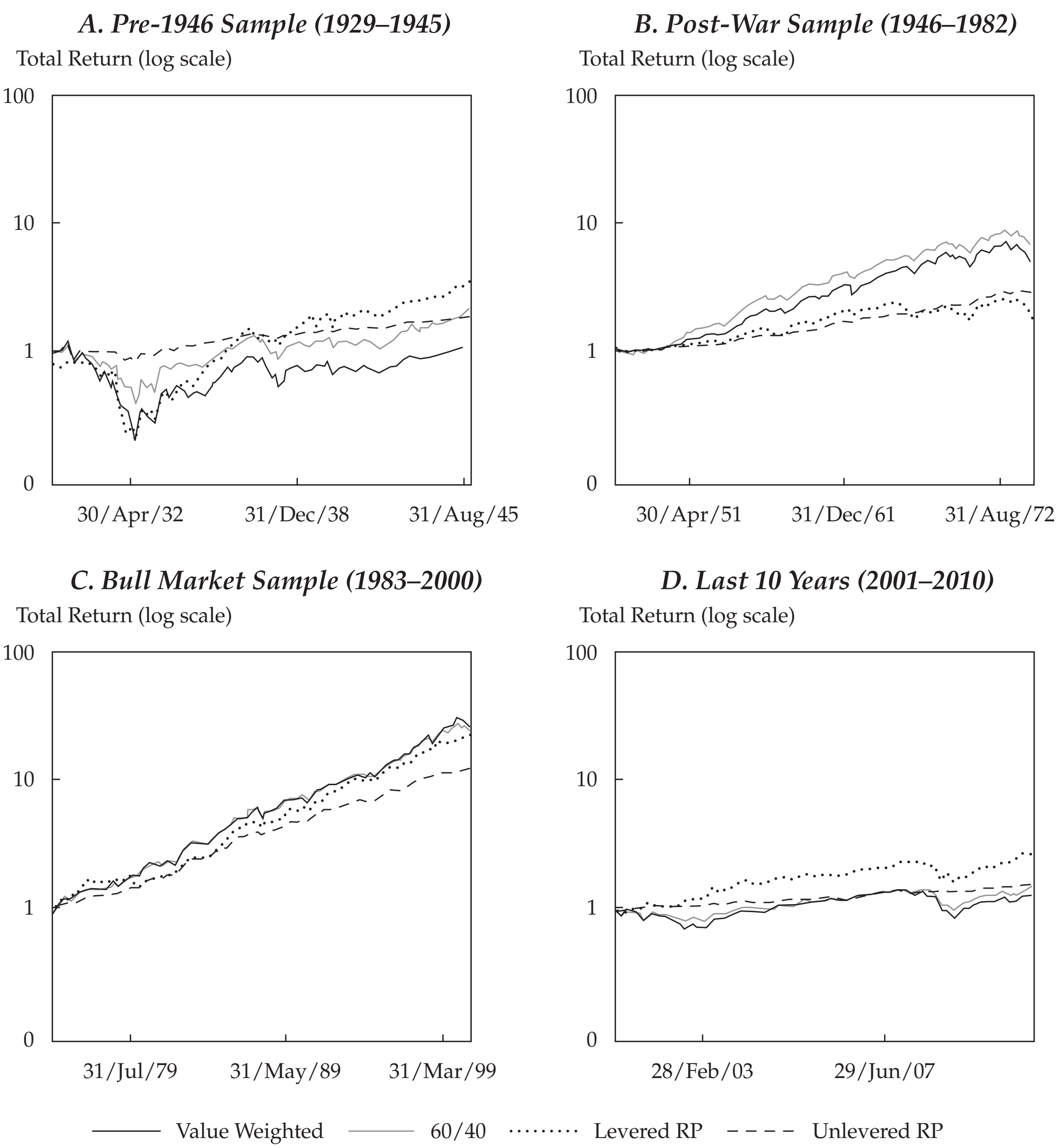
changes.[12] As suggested by **Figure 5**, leverage exacerbates turnover, and so the trading costs for the levered risk parity strategy are much higher than those for the unlevered risk parity and 60/40 strategies. However, the data required to determine the precise relationship between turnover and trading costs were unavailable. So, we estimated.[13]

Figure 6 shows the cumulative return to the four strategies over the long horizon. We financed the levered risk parity strategy at the three-month Eurodollar deposit rate. We incorporated the turnover-induced trading costs into the returns to the 60/40 and risk parity strategies. With respect to cumulative 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.

Statistical Significance of Findings

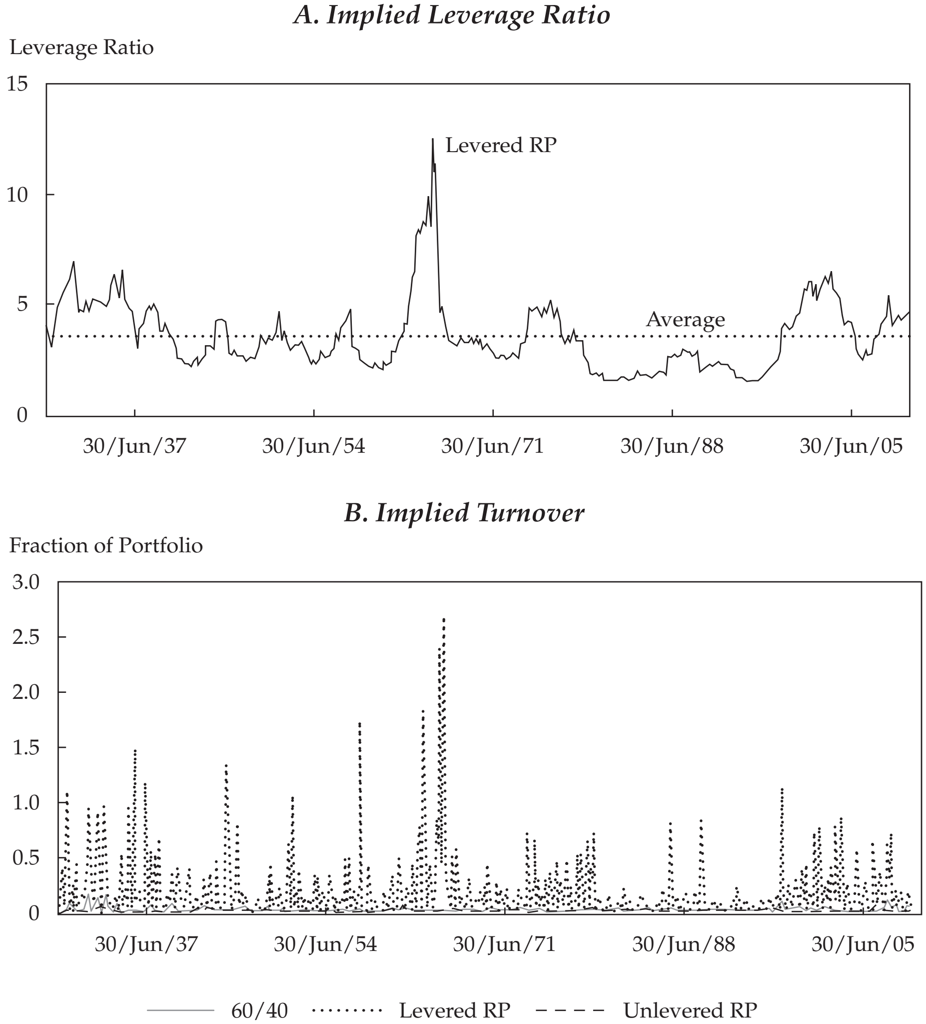
Because the volatility of an asset's return is substantially greater than its expected value, achieving statistical significance in a comparison of investment strategies is difficult, even over periods of decades. **Table 1** presents p-values[14] for these comparisons. Disregarding trading costs and assum-

Figure 4. Cumulative Returns for Risk Parity vs. the Market vs. 60/40 over Various Sample Periods



Notes: This figure shows monthly compounded returns to four strategies based on U.S. equity and U.S. Treasury bonds over four subperiods. The levered risk parity strategy was financed at the three-month Eurodollar deposit rate. A comparison with Figure 2 shows the magnitude of the performance drag, which was most severe in the post-war sample.

Figure 5. Implied Leverage Ratio and Implied Turnover, 1929–2010



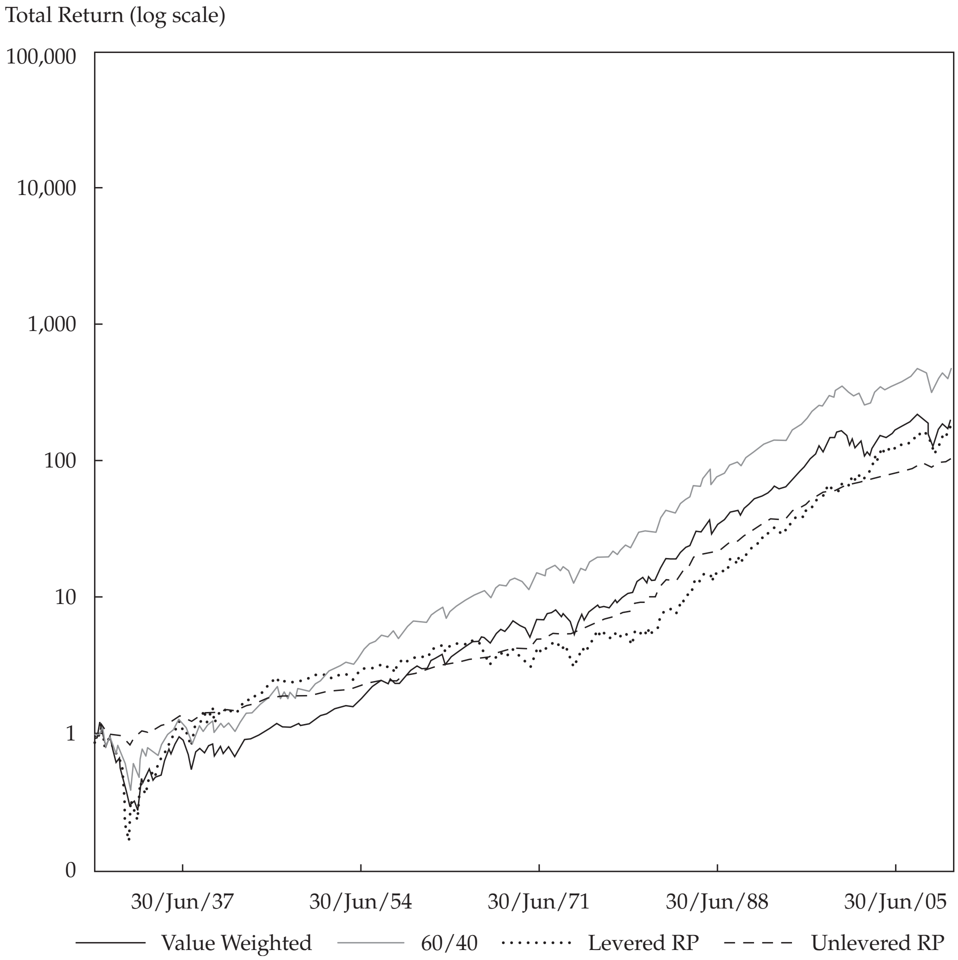
Notes: Panel A plots the leverage required for the estimated volatility of the risk parity strategy to match the estimated volatility of the market at each rebalancing. The average over the entire period is 3.55. The spike in leverage occurred on 30 September 1965, which was a rare moment when bond volatility was relatively low (0.5%) and both equity volatility (10%) and market weight (72%) were relatively high. Panel B shows the turnover of the risk parity and 60/40 strategies at each rebalancing.

ing borrowing at the risk-free rate, the (annualized monthly arithmetic) mean return of levered risk parity exceeded that of 60/40 in the 85-year long sample by 210 bps, and the result is statistically significant (p = 0.03). However, 60/40 was somewhat less volatile than levered risk parity; taking this fact into account, the alpha for levered risk parity minus 60/40 just misses being significant (p = 0.06).

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

Let us turn the problem around. Suppose we ignore trading costs and assume we can borrow at the risk-free rate. Suppose that, on the basis of 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 bps. 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 17.5%. So, if we ignore borrowing and trading costs, 60/40 has a substantial probability of beating levered risk parity over the next 20 years and even over the next 50 years.

Figure 6. Cumulative Returns for Risk Parity vs. the Market vs. 60/40, 1929–2010



Notes: This figure shows monthly compounded returns to four strategies based on U.S. equity and U.S. Treasury bonds. The levered risk parity strategy was financed at the three-month Eurodollar deposit rate, and adjustments were made for turnover. A comparison with Figure 3 shows the magnitude of the performance drag.

Of course, if we do take into account 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.

Risk Profiles

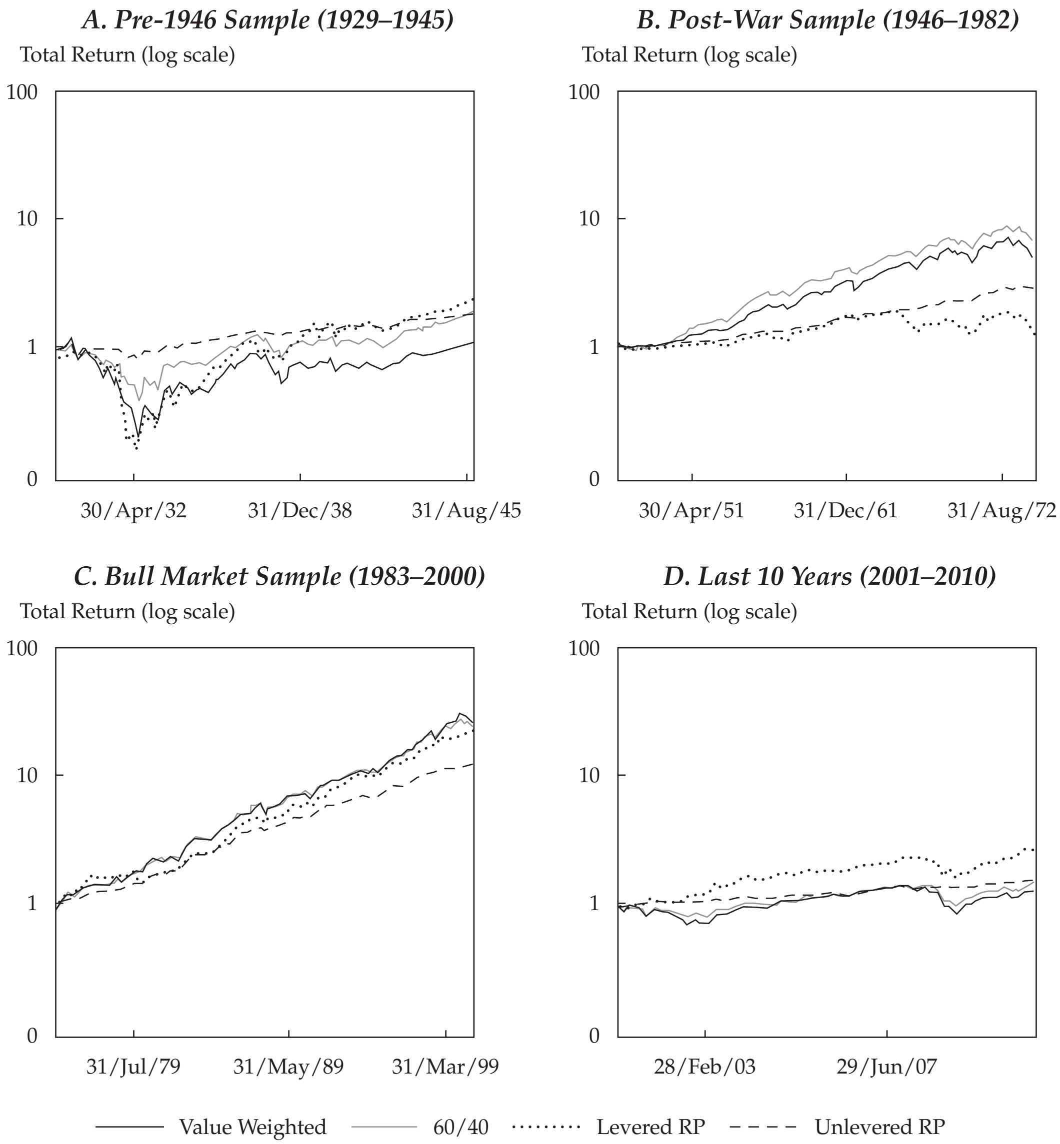
Because a thorough evaluation of investment strategies involves risk as well as return, we considered the realized Sharpe ratios of the four strategies. **Figure 8** shows the strategy Sharpe ratios over 1926–2010, and the subperiod Sharpe ratios are shown in **Figure 9**. Both figures indicate that *unlevered risk parity* had the highest realized Sharpe ratio, with 60/40 coming in second.[16] In the CAPM, the value-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. This type of weighted combination maximizes return for given levels of risk. However, in the more realistic case where 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

Figure 7. Cumulative Returns for Risk Parity vs. the Market vs. 60/40 over Various Sample Periods



Notes: This figure shows monthly compounded returns to four strategies based on U.S. equity and U.S. Treasury bonds over four subperiods. The levered risk parity strategy was financed at the three-month Eurodollar deposit rate, and adjustments were made for turnover. A comparison with Figure 4 shows the magnitude of the performance drag.

is the volatility of the unlevered portfolio. For large leverage,

$$S_L \approx S_U - \left(\frac{r_b - r_f}{\sigma}\right). \quad (2)$$

When the borrowing rate exceeds the risk-free rate, the efficient frontier comprises 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 Equation 1, is equal to the slope of the line connecting the portfolio to the risk-free portfolio.

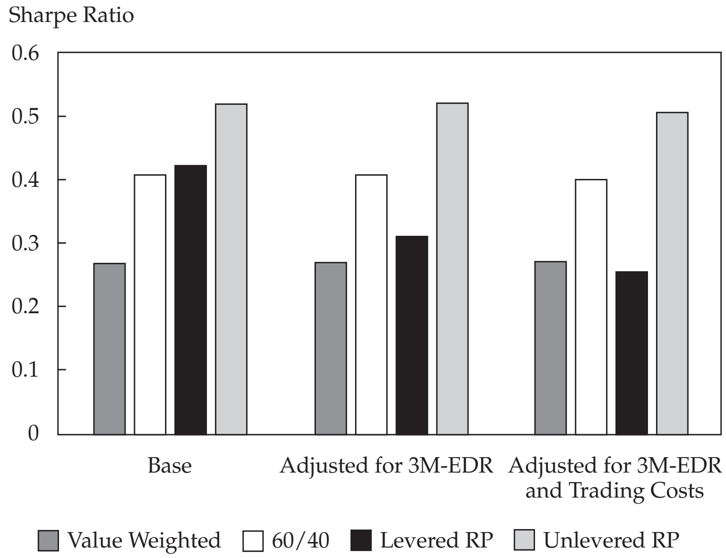
Why did the levered risk parity strategy in Asness, Frazzini, and Pedersen (2012) outperform the others after adjusting for financing costs in excess of the risk-free rate, whereas an analogous adjustment to our levered risk parity strategy caused it to underperform? They matched the long sample *ex post* volatility of the levered risk parity

Table 1. Risk Parity vs. the Market vs. 60/40: Historical Performance, 1926–2010

Base Case	Excess Return	*p*-Value Excess Return	Alpha	*p*-Value Alpha	Volatility	Sharpe Ratio	Skewness	Excess Kurtosis
A. Long sample stocks and bonds, 1926–2010								
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) – value weighted	2.84	0.01	3.53	0.00	10.73	0.26	–0.51	12.42
Risk parity (levered) – 60/40	2.10	0.03	1.81	0.06	10.11	0.21	–1.08	13.58
B. Long sample stocks and bonds, 1926–2010, adjusted for 3M-EDR								
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) – value weighted	1.03	0.20	1.70	0.07	10.72	0.10	–0.57	12.50
Risk parity (levered) – 60/40	0.29	0.40	–0.02	0.50	10.11	0.03	–1.15	13.68
C. Long sample stocks and bonds, 1926–2010, adjusted for 3M-EDR and trading costs								
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) – value weighted	0.11	0.47	0.79	0.24	10.75	0.01	–0.67	13.06
Risk parity (levered) – 60/40	–0.51	0.67	–0.81	0.77	10.13	–0.05	–1.22	13.98

Notes: This table reports performance statistics for the four strategies and CRSP stocks and bonds over 1926–2010. In Panel A, the levered risk parity strategy was financed at the 90-day T-bill rate. In Panels B and C, the levered risk parity strategy was financed at the three-month Eurodollar deposit rate (3M-EDR). In Panel C, the 60/40 and risk parity strategies were adjusted for turnover.

Figure 8. Sharpe Ratios for Risk Parity vs. the Market vs. 60/40, 1929–2010



Notes: This figure shows the realized Sharpe ratios for the four strategies; 3M-EDR stands for the three-month Eurodollar deposit rate. Unlevered risk parity dominates even before adjustment for market frictions.

strategy with the long sample *ex post* volatility of the value-weighted strategy. Of course, this volatility cannot be known in advance, and so their levered risk parity strategy is not investable.

Table 1 also reports standard statistics for the four strategies. The best-performing strategy depends on how an investor weights various risk and performance measures. For example, when one has positive skewness, high kurtosis may be desirable, and this combination occurs for the value-weighted and 60/40 strategies. Levered risk parity exhibits negative skewness and high kurtosis. This bad combination can lead to deleveraging costs, which can 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 explanation for the performance of the frictionless version of levered risk parity: Perhaps there is a premium for taking on severe downside risk.[18]

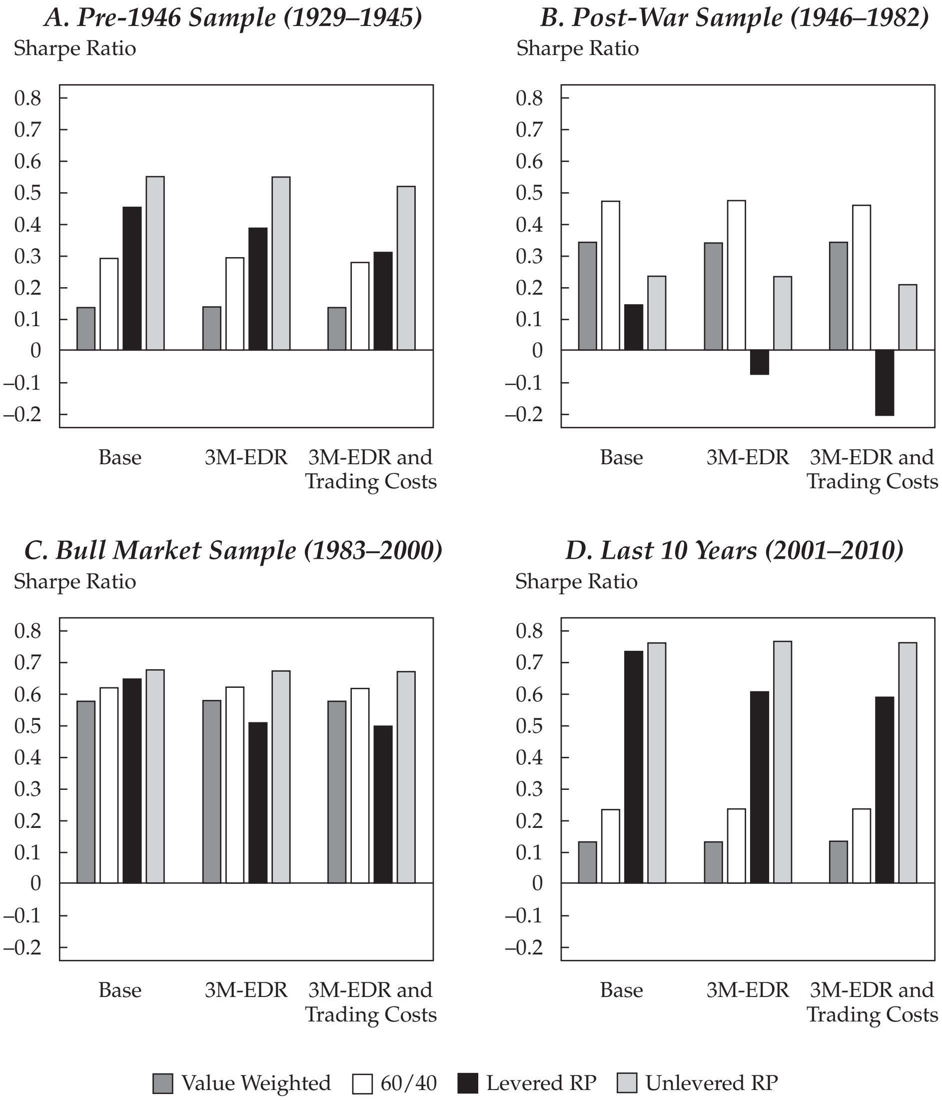
Conclusion

Strategy evaluation is an important part of the investment process. However, because 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 various lengths and in various 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 that 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 that we implemented returned substantially more than did unlevered risk parity, a 60/40 fixed mix, and a value-weighted portfolio. There are, however, important caveats. First, levered risk parity underperformed over 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. Unlevered risk parity dominated our study on the basis of risk-adjusted return, or realized Sharpe ratio. Other performance measures might lead to different conclusions.

Compelling economic theories of leverage aversion give credence to the idea that levered risk parity may outperform the market over long horizons (see, e.g., Frazzini and Pedersen 2010). However, there are dissenting voices, which are also compelling (see, e.g., Sullivan 2010). Our findings suggest that risk parity may be a preferred strategy under

Figure 9. Sharpe Ratios for Risk Parity vs. the Market vs. 60/40 over Various Sample Periods



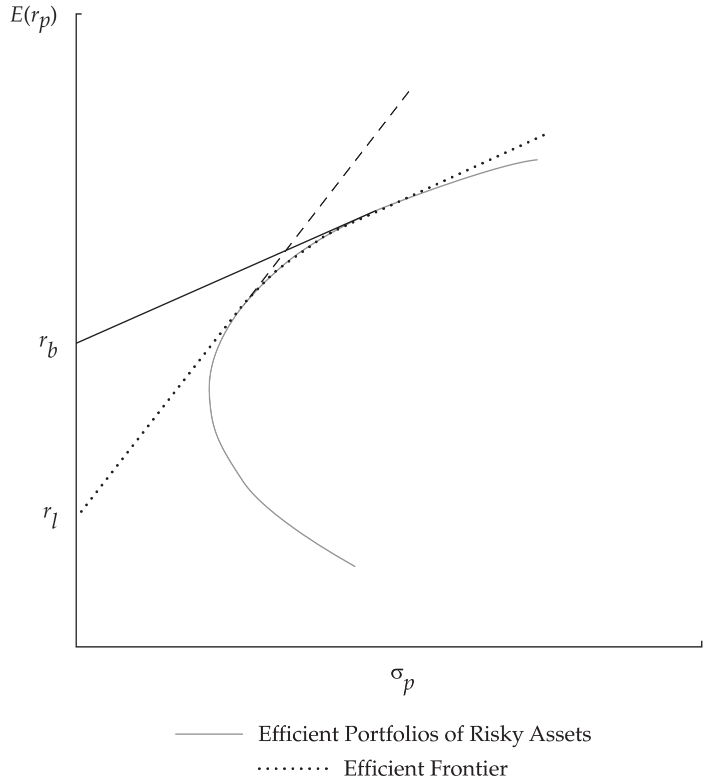
Notes: This figure shows the realized Sharpe ratios for the four strategies over four subperiods; 3M-EDR indicates Sharpe ratios adjusted for the three-month Eurodollar deposit rate, and 3M-EDR and Trading Costs indicates Sharpe ratios adjusted for the three-month Eurodollar deposit rate and trading costs. Apart from the post-war sample, unlevered risk parity dominates even before adjustment for market frictions.

certain market conditions or with respect to certain yardsticks. But any inference from our results must take into account 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.

We thank Bernd Scherer for careful reviews and insightful comments that improved this article; we also thank Günter Schwarz for educating us about the history of risk parity and for providing us with early references on the subject. We gratefully acknowledge the valuable feedback on our ideas and results from the participants in the Q Group Spring 2012 Seminar. Robert Anderson and Stephen Bianchi were supported by the Coleman Fung Chair in Risk Management at the University of California, Berkeley.

This article qualifies for 1 CE credit.

Figure 10. Sharpe Ratio of a Levered Portfolio on the Efficient Frontier



Notes: 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.

Appendix A. Data, Strategies, Trading Costs, Bootstrap Estimates, and Robustness Checks

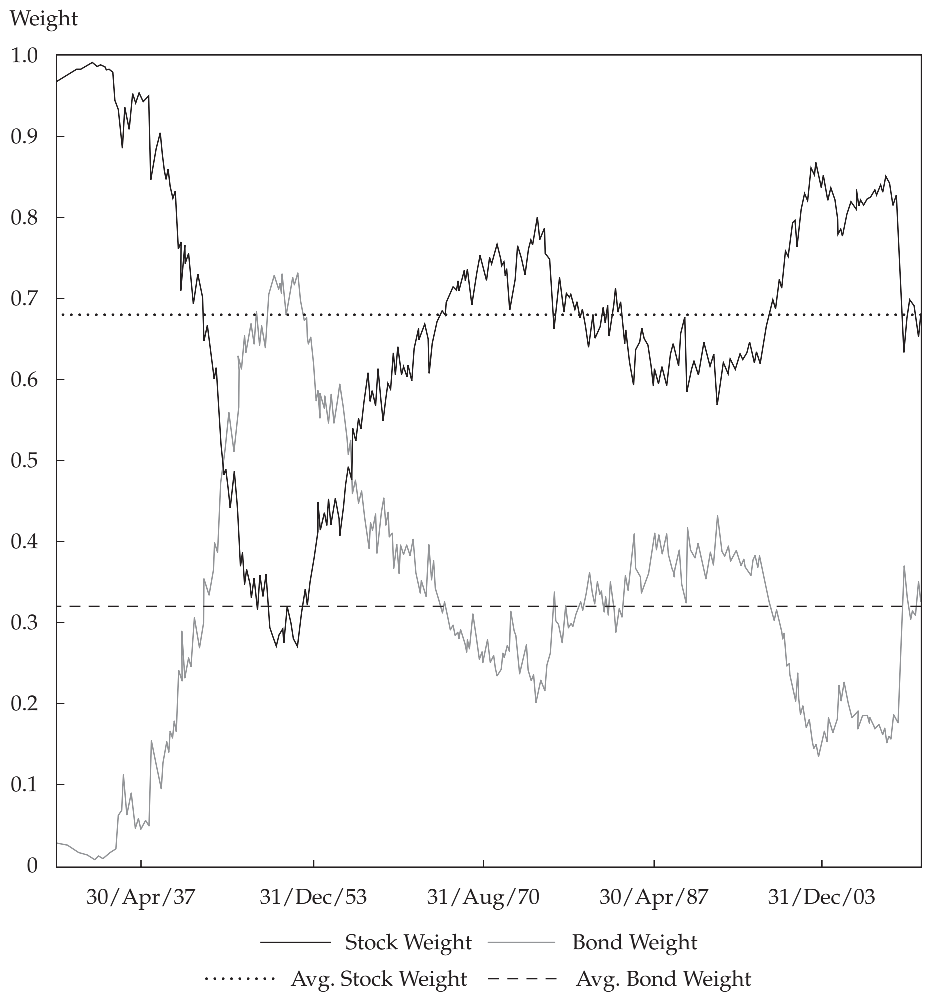
This appendix provides details on our data sources and definitions, strategies used, trading cost and bootstrap estimates, and robustness checks with respect to the rebalancing horizon and borrowing cost assumptions.

Data

The results presented in this article are based on CRSP stock and bond data for January 1926–December 2010. The aggregate stock return is the CRSP value-weighted market return (including dividends) from the Monthly Stock Market Indices (NYSE/AMEX/NASDAQ) table (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 which the variable name for the unadjusted return is *retnua* and for the face value outstanding is *iout1r*. We used all bonds in the table, provided that the values for both *retnua* and *iout1r* were not missing. We constructed the value-weighted market index 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 A1** 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 U.S. government 90-day T-bill secondary market rate, provided by Global Financial Data, covering January 1926–December 2010. The proxy for the cost of financing leverage is the U.S. three-month Eurodollar deposit rate, obtained from the Federal Reserve's daily selected interest rates webpage (www.federalreserve.gov/releases/h15/data.htm). The three-month Eurodollar deposit data are available for January 1971–December 2010. Prior to January 1971, a constant of 60 bps is added to the 90-day T-bill rate.[19]

Figure A1. Stock and Bond Weights in the Value-Weighted Index, 1929–2010



Note: This figure shows the weights for stocks and bonds implied by market capitalization.

Strategies

We rebalanced the strategies monthly.

Value weighted. This fully invested strategy value weights U.S. equity and U.S. Treasury bonds.

60/40. This fully invested strategy's capital allocations are 60% U.S. equity and 40% U.S. Treasury bonds.

Unlevered risk parity. This fully invested strategy equalizes *ex ante* asset class volatilities. The volatility of each asset class is estimated at month end by using a 36-month rolling window of trailing returns. The time t estimate of volatility for asset class i is given by

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

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

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

where

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

Levered risk parity. This levered strategy 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, the volatility of a strategy is estimated at month end by using a 36-month rolling window of trailing returns. The time t estimate of volatility for strategy s is given by

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

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, $\hat{\sigma}_{v,t}$, and 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

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

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

Asness, Frazzini, and Pedersen (2012) implemented an unconditional levered risk parity strategy. The asset class weights in this strategy depend on a time-independent scale factor k, chosen so that the 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,

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

$$r_{l,unc,t}^{e} = r_{l,unc} - r_{f,t},$$

and

$$\sigma = \text{std}\left(r_{l,unc,37}^{e}, \ldots, r_{l,unc,T}^{e}\right), \tag{A1}$$

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–December 2010, then T = 1,020). Note that the target σ is not known until the end of the period. Moreover, even if σ were set to some constant that was 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 A1 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 being equal, both strategies will increase the ratio of capital in bonds to capital in equity, but the conditional strategy will increase its leverage, whereas the unconditional strategy will decrease its leverage.

Trading Costs

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

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

and the turnover required to rebalance the strategy is given by

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

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

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

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. Trading cost–adjusted returns are given by

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

Bootstrap Estimates

To reflect the empirical properties of our data, we used a bootstrap to estimate the p-values in Table 1. For a given strategy and evaluation period, let us 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 that

$$\mathbf{R}_s = \alpha + \beta\mathbf{R}_b + \varepsilon, \tag{A2}$$

where $\mathbf{R}_b$ is the vector of excess returns of a benchmark portfolio—that is, $\mathbf{R}_b = (R_{b,1}, \ldots, R_{b,T})'$—which, in our case, is the value-weighted portfolio, and $\mathbf{R}_s$ 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

$$\mathbf{R}_s^* = \hat{\alpha} + \hat{\beta}\mathbf{R}_b + \varepsilon^*,$$

where ε^* is the vector of resampled residuals. Then, for each sample, we run a time-series regression based on Equation A2 to obtain new estimates of alpha $\left(\hat{\alpha}^*\right)$ and beta $\left(\hat{\beta}^*\right)$. The p-value for alpha is given by

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

The probability estimates in the article 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 the 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

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

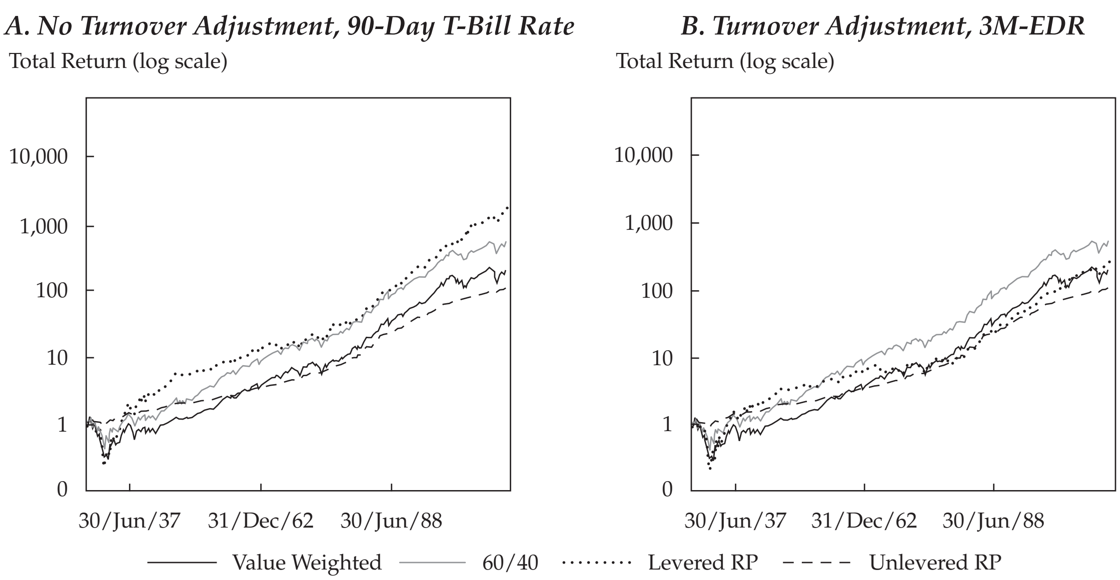
Two Robustness Checks

We considered the impact of two of the assumptions that we made in our study.

Rebalancing horizon. The monthly rebalancing horizon that we used is shorter than the horizon typically used by pensions, endowments, and other long-term investors. **Figure A2** shows the cumulative returns to the four strategies when they are rebalanced annually. Panel A shows the base case in which borrowing is at the 90-day T-bill rate and no adjustment is made for trading costs arising from turnover. In Panel B, borrowing is at the three-month Eurodollar deposit rate and adjustments are made for trading costs arising from turnover. The results are less dramatic than they are when the horizon is rebalanced monthly, but they are qualitatively similar.

Borrowing cost assumptions. We used the U.S. three-month Eurodollar deposit rate, starting in 1971, as the estimate of the implicit interest rate when levering with futures. Over 1926–1970, when the three-month Eurodollar deposit rate was unavailable, we extrapolated borrowing costs to be the 90-day T-bill rate plus 60 bps—a conservative estimate because the average spread between the three-month Eurodollar deposit rate and the 90-day T-bill rate over 1971–2010 is roughly 100 bps. **Table A1** shows the impact of varying the extrapolated

Figure A2. Cumulative Returns to the Four Strategies When Rebalanced Annually, 1929–2010



Notes: This figure shows continuously compounded returns to the four strategies with annual rebalancing. In Panel A, the levered risk parity strategy was financed at the 90-day T-bill rate and no adjustments were made for turnover. In Panel B, the levered risk parity strategy was financed at the three-month Eurodollar deposit rate and adjustments were made for turnover.

spread on strategy performance. We considered spreads ranging between 25 and 125 bps. 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 bps was assumed to be the spread between the three-month Eurodollar deposit rate and the 90-day T-bill rate over 1926–1970.

Table A1. Effect of Alternative Borrowing Cost Assumptions for the Pre-1971 Period

Pre-1971 Borrowing Cost Spread over T-Bills	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 bps	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 bps	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 bps	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 bps	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

Notes: This table reports the impact of the borrowing cost extrapolation on the risk premium of levered risk parity over the value-weighted strategy (1926–2010). In Panels A and B, the levered risk parity strategy was financed at the three-month Eurodollar deposit rate (3M-EDR). In Panels C and D, an additional adjustment was made for trading costs arising from turnover.

Notes

1. Our simple two-asset-class strategies, which involve neither market timing nor security selection, can be used as benchmarks to evaluate more complex strategies.
2. Lörtscher (1990) and Kessler and Schwarz (1996) used the term *equal risk benchmarks* for risk parity strategies.
3. For a discussion of the long-term outlook for risk parity strategies, see Veronica Dagher, "New Allocation Funds Redefine Idea of 'Balance,'" *Wall Street Journal* (6 February 2012).
4. For example, Qian (2005) examined the implications of including asset correlations in risk parity weights. Chaves, Hsu, Li, and Shakernia (2011) considered a broader array of asset classes and also considered risk parity in the context of other low-risk strategies.
5. There is a large and growing literature on low-risk investing. Sefton, Jessop, De Rossi, Jones, and Zhang (2011) gave a broad discussion of the topic, and Scherer (2011) attributed to Fama–French factors the empirically observed outperformance of the market by a particular low-risk (minimum variance) strategy. Cowan and Wilderman (2011) provided a rational explanation for the low-risk anomaly, and Baker, Bradley, and Wurgler (2011) offered a behavioral explanation. Clarke, de Silva, and Thorley (2011) analyzed the connection between low-beta and minimum variance strategies.
6. In our study, we used the 90-day T-bill rate as a proxy for the risk-free rate.
7. For such a liquid asset class as U.S. Treasury bonds, futures may be the cheapest way to finance the levered position. However, U.S. Treasury futures have been traded in a liquid market only since the 1980s. Therefore, it is impossible to conduct a fully empirical study of risk parity that begins early in the 20th century because we cannot know how a futures-financed risk parity strategy would have performed during the Great Depression. Although we can estimate what it would have cost to finance the leverage through more conventional borrowing, small differences in assumptions about borrowing costs can have major effects on the return estimates for a levered risk parity strategy, precisely because the strategy involves such a high degree of leverage. Moreover, because the introduction of liquid U.S. Treasury futures markets presumably reduced the cost of financing a levered risk parity strategy, it may have induced changes in asset returns that would tend to offset the savings achieved through lower financing costs.
8. Specifically, borrowing was at the three-month Eurodollar deposit rate starting in 1971 and was equal to the risk-free rate plus 60 bps before 1971. Turnover-induced trading costs were 1% over 1926–1955, 0.5% over 1956–1970, and 0.1% over 1971–2010.
9. Chaves, Hsu, Li, and Shakernia (2011, p. 111) commented 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."
10. Asness, Frazzini, and Pedersen (2012) found that their unconditional levered risk parity, when financed at LIBOR, outperformed both 60/40 and value-weighted strategies over the long sample. They asserted 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" (p. 57). We found 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 account. We also found 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. Note that although many of the figures in this article show results for 1929–2010, our data sample runs from 1926 through 2010. We used a three-year window to estimate volatility, and thus, the results for any strategy that depends on a volatility estimate begin in 1929.
11. Over the period when both the U.S. three-month Eurodollar deposit rate and the three-month LIBOR are available, they track each other very closely, with LIBOR being about 10 bps higher, on average.
12. The details of our turnover estimates are in Appendix A.
13. We assumed trading costs of 1% over 1926–1955, 0.5% over 1956–1970, and 0.1% over 1971–2010.
14. In 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.
15. There is an apparent conflict between the information in Panel B of Table 1 and the information in Figure 3. Table 1 shows that if we take into account borrowing costs that exceed the risk-free rate but do not adjust for trading costs arising from turnover, levered risk parity outperforms 60/40 by a (statistically insignificant) 29 bps. Under the same assumptions, Figure 3 shows that 60/40 outperforms levered risk parity over 1926–2010. Table 1 reports the arithmetic mean of the monthly returns, which does not handle compounding correctly. Figure 3 presents the cumulative returns for 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, 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 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).
17. Markowitz (2005) discussed a simple paradigm in which leverage constraints render the market portfolio inefficient in an idealized setting.
18. We thank an anonymous referee for the observations in this paragraph.
19. The average difference between the 90-day T-bill rate and the three-month Eurodollar deposit rate over 1971–2010 is roughly 100 bps. Thus, our estimate of 60 bps is relatively conservative.

References

Asness, Clifford S., Andrea Frazzini, and Lasse H. Pedersen. 2012. "Leverage Aversion and Risk Parity." *Financial Analysts Journal*, vol. 68, no. 1 (January/February):47–59.

Baker, Malcolm, Brendan Bradley, and Jeffrey Wurgler. 2011. "Benchmarks as Limits to Arbitrage: Understanding the Low-Volatility Anomaly." *Financial Analysts Journal*, vol. 67, no. 1 (January/February):40–54.

Black, Fischer. 1972. "Capital Market Equilibrium with Restricted Borrowing." *Journal of Business*, vol. 45, no. 3 (July):444–455.

Black, Fischer, Michael C. Jensen, and Myron Scholes. 1972. "The Capital Asset Pricing Model: Some Empirical Tests." In *Studies in the Theory of Capital Markets*. Edited by Michael C. Jensen. New York: Praeger Publishers.

Chaves, Denis, Jason Hsu, Feifei Li, and Omid Shakernia. 2011. "Risk Parity Portfolio vs. Other Asset Allocation Heuristic Portfolios." *Journal of Investing*, vol. 20, no. 1 (Spring):108–118.

Clarke, Roger, Harindra de Silva, and Steven Thorley. 2011. "Minimum-Variance Portfolio Composition." *Journal of Portfolio Management*, vol. 37, no. 2 (Winter):31–45.

Cowan, David, and Sam Wilderman. 2011. "Rethinking Risk: What the Beta Puzzle Tells Us about Investing." White paper, GMO (November).

Frazzini, Andrea, and Lasse H. Pedersen. 2010. "Betting against Beta." NBER Working Paper 16601 (December).

Kessler, Christoph, and Günter Schwarz. 1996. "Investment Management for Private Investors at SBC—The Benchmark." Research report (in *The Global*), Swiss Bank Corporation (2nd Quarter):1–12.

Lörtscher, Rudolf. 1990. "Investment Policy at SBC—Basic Premises and Model." Research report (in *The Global*), Swiss Bank Corporation (October):20A–20D.

Markowitz, Harry M. 2005. "Market Efficiency: A Theoretical Distinction and So What?" *Financial Analysts Journal*, vol. 61, no. 5 (September/October):17–30.

Naranjo, Lorenzo. 2009. "Implied Interest Rates in a Market with Frictions." Working paper, ESSEC Business School (February).

Qian, Edward. 2005. "Risk Parity Portfolios: Efficient Portfolios through True Diversification." Research report, PanAgora (September).

Scherer, Bernd. 2011. "A Note on the Returns from Minimum Variance Investing." *Journal of Empirical Finance*, vol. 18, no. 4 (September):652–660.

Sefton, James, David Jessop, Giuliano De Rossi, Claire Jones, and Heran Zhang. 2011. "Q-Series: Low-Risk Investing." Research report, UBS Investment Research (23 September).

Sullivan, Rodney N. 2010. "Speculative Leverage: A False Cure for Pension Woes." *Financial Analysts Journal*, vol. 66, no. 3 (May/June):6–8.