Will My Risk Parity Strategy Outperform?

Robert M. Anderson*
University of California at Berkeley
Stephen W. Bianchi†
University of California at Berkeley
Lisa R. Goldberg‡
MSCI and University of California at Berkeley

November 10, 2011§

Abstract

We gauge the return-generating potential and risk inherent in four investment strategies: value weighted, fixed mix, and levered and unlevered risk parity, over an 85-year horizon. There are three essential conclusions from our study. First, even over periods lasting decades, the specific start and end dates of a backtest can have a material effect on the results; second, transaction costs can negate apparent outperformance; third, statistical significance of findings needs to be assessed.

Key words: Risk parity, value weighted, leverage, turnover, trading costs, market frictions, statistical significance

*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: swbianchi@berkeley.edu.

‡2100 Milvia Street, Berkeley, CA 94704-1113, USA, email: lisa.goldberg@msci.com.

§We thank Rodney Sullivan for stimulating conversations about risk parity, leverage and strategy evaluation.

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 material effect on measured returns. In addition, the introduction of new securities can have a material effect on the returns of a strategy; a strategy that seems to have been profitable in the past may become less profitable if the introduction of new securities makes it easier for others to replicate the strategy than was the case in the past.

We consider these issues in the context of a risk parity strategy, which equalizes risk contributions across asset classes, and is typically levered to match market levels of risk. The cost of financing the high degree of leverage in a levered risk parity strategy can have a material impact on the performance of a strategy. 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 levered risk parity, 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 probably induced changes in asset returns that would tend to offset the savings achieved through lower financing costs.

Even assuming that the underlying processes possess some strong form of stationarity, the high volatility of security returns poses two closely related 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.

In this article, we examine the historical performance of four strategies based on two asset classes: US Equity and US Treasury Bonds. Our study includes a value weighted portfolio, a 60/40 mix, and two risk parity strategies. 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 market,[1] we obtain levered risk parity. The strategies are rebalanced each month.

Strategy performance is evaluated in different time periods and before and after adjusting for market frictions. Our long study period is 1926–2010, and we also consider four important subperiods. The Pre-1946 Sample, 1926–1945, which included the Great Depression and World War II, was also plagued by deflationary shocks and inflationary spikes. From an historical perspective, equity markets were relatively calm during the The Post-War Sample, 1946–1982. However, this period included a bout of severe inflation and high interest rates that translated into poor stock and bond performance. The Bull Market Sample, 1983–2000, included a huge bond rally and the game-changing emergence of the technology industry. The Last 10 Years felt turbulent, although they were, in fact, much calmer than the initial years of the study period. The information required to reproduce our results is in Appendices A and B.

We find that strategy performance depends materially on the analysis period. Our results are consistent with, but not sufficient to demonstrate, the assertion that risk parity tends to outperform in turbulent markets. By extrapolating transaction costs based on recent experience to our entire study period, we find market frictions are a substantial drag on performance. However, since we do not know how the availability of modern financing methods during the period 1926–1971 might have affected the course of history, our results should be interpreted with caution.

2 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. In the most recent period, even unlevered risk parity beat the value weighted and 60/40 strategies. 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 strategies tied for first place.

[1] Throughout this article, we consider a simplified market consisting only of US Equity and US Treasury Bonds weighted according to market value.

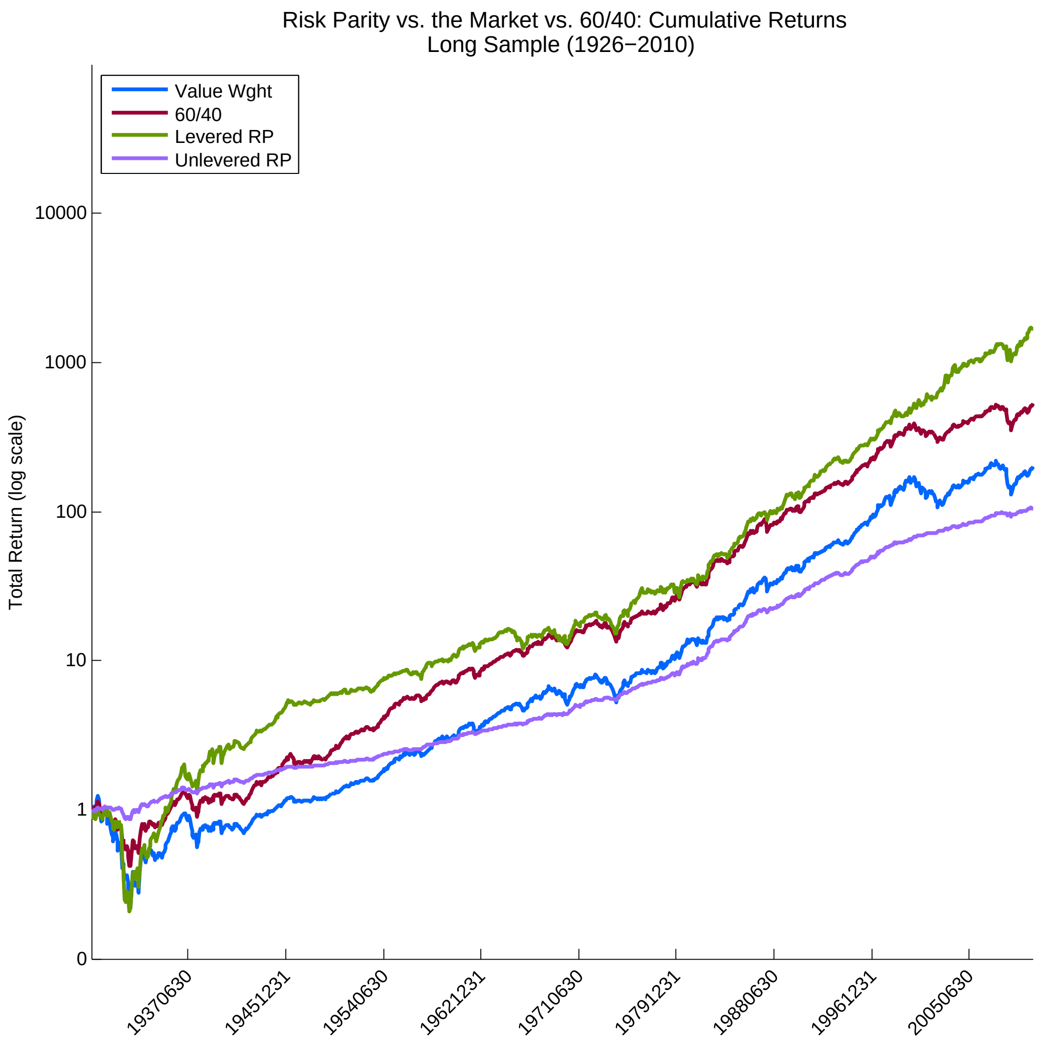

Figure 1: Continuously compounded return to four strategies based on US Equity and US Treasury Bonds over the period 1926–2010. The levered risk parity strategy is financed at the 90-Day T-Bill rate.

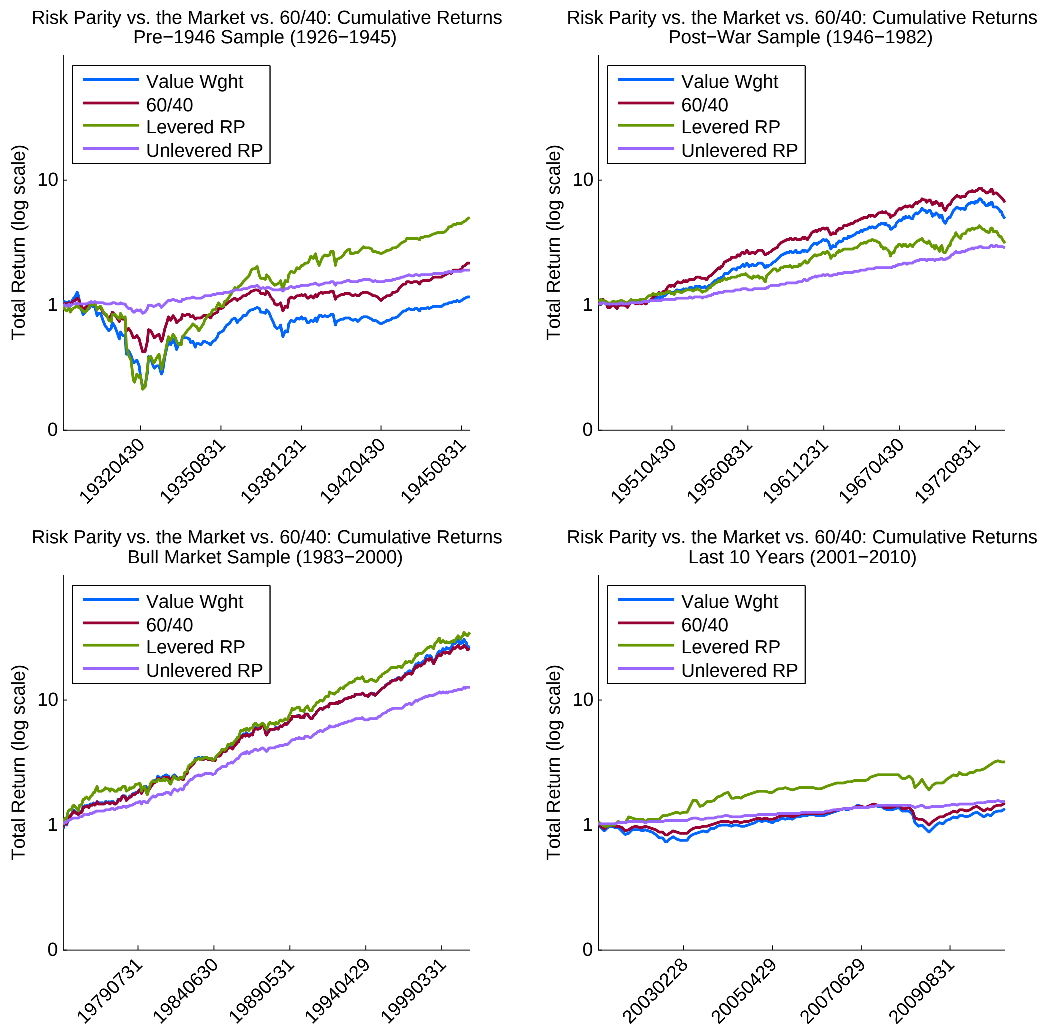

Figure 2: Continuously compounded return to four strategies based on US Equity and US Treasury Bonds over 4 subperiods. The levered risk parity strategy is financed at the at the 90-Day T-Bill rate. The results depend materially on the evaluation period.

3 Transaction Costs Can Negate Apparent Outperformance

3.1 Borrowing Costs

In the studies discussed in Section 2, we a 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 rate at which investors can actually borrow. Because it is available over a longer period, we use the US 3-Month Euro-Dollar Deposit Rate as a proxy for LIBOR.[2] We repeat the studies in Section 2 replacing the 90-day T-Bill rate with the 3-Month Euro-Dollar Deposit 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.

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. 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.

3.2 Trading Costs

Value weighted strategies require rebalancing only in response to a limited set of events: 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 and 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.[3] 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.[4]

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

[2]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.

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

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

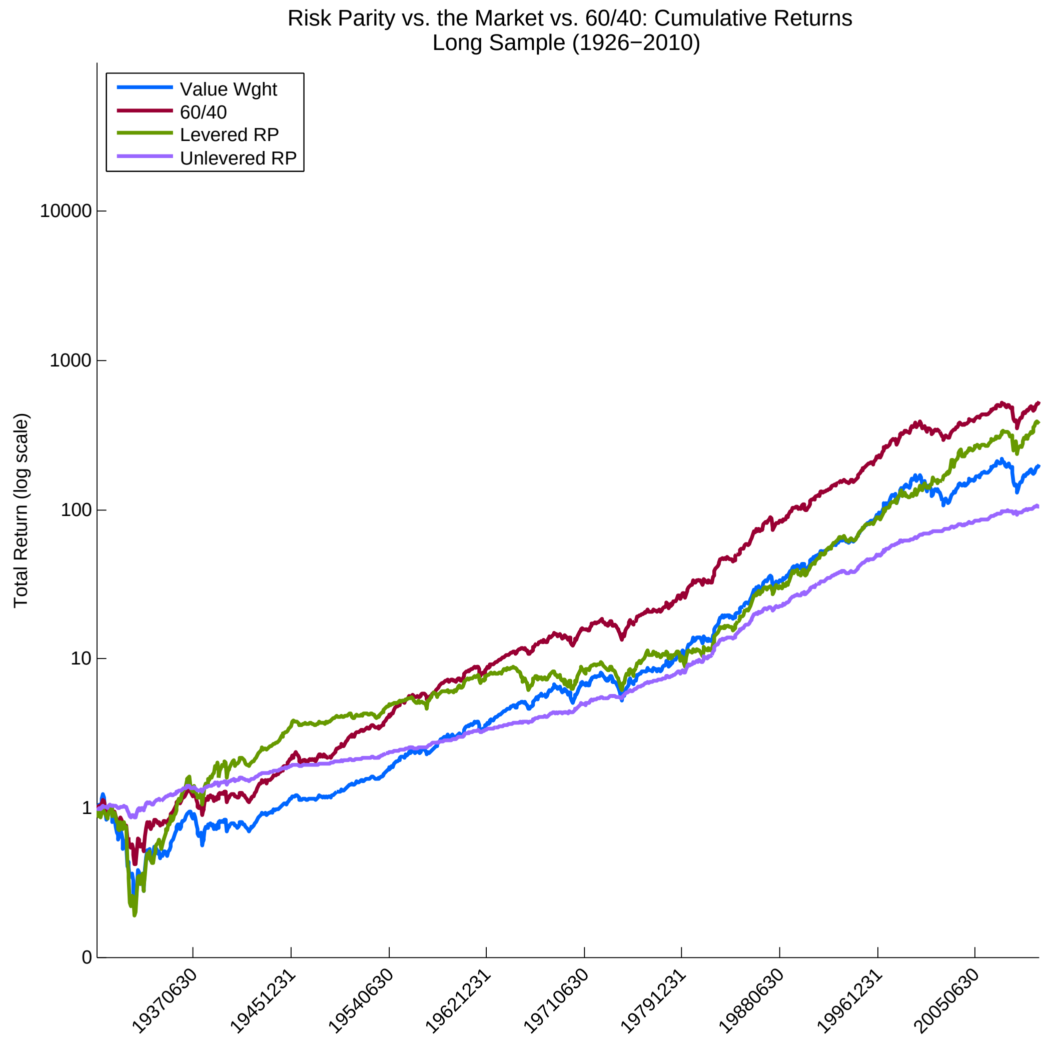

Figure 3: Continuously compounded return to four strategies based on US Equity and US Treasury Bonds over the period 1926–2010. The levered risk parity strategy is financed at the a 3-Month Euro-Dollar Deposit rate. A comparison with Figure 1 shows the magnitude of the performance drag.

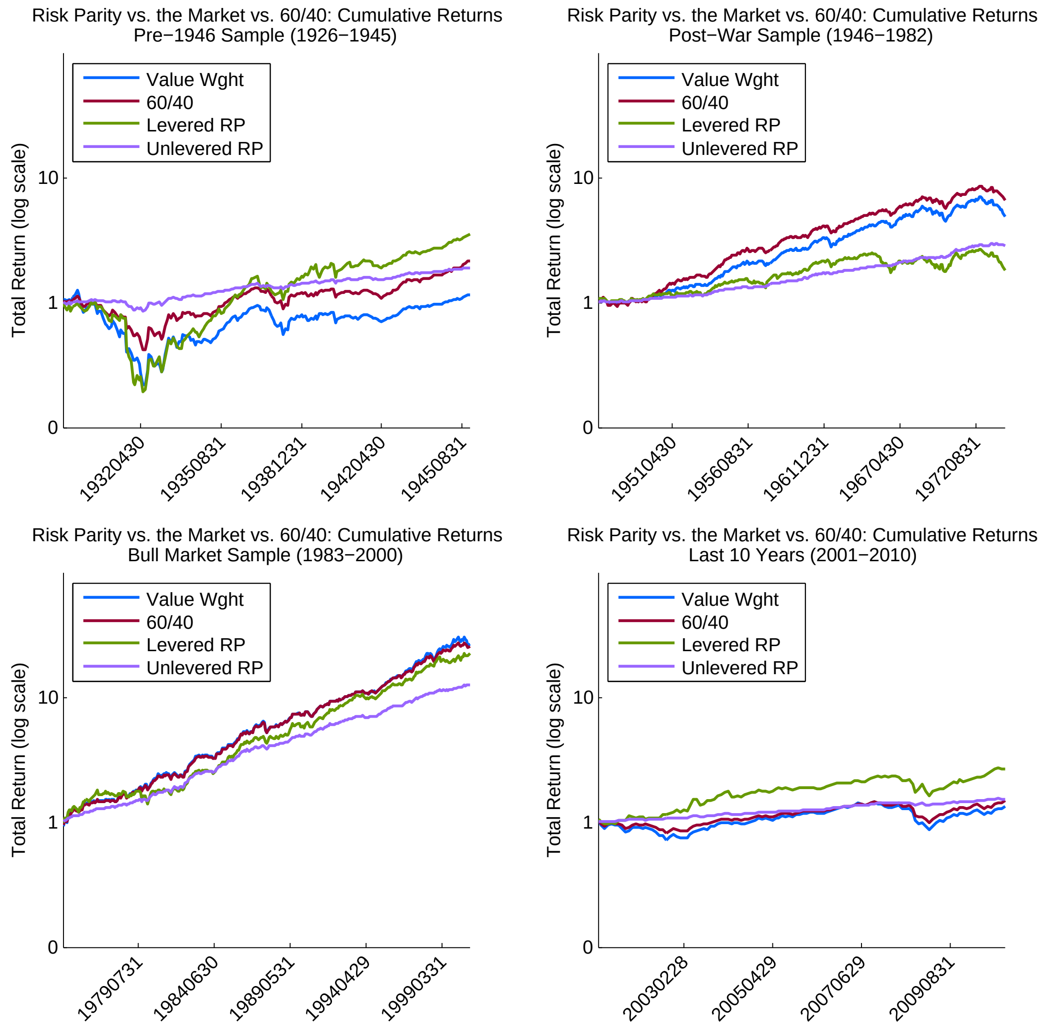

Figure 4: Continuously compounded return to four strategies based on US Equity and US Treasury Bonds over 4 subperiods. The levered risk parity strategy is financed at the a 3-Month Euro-Dollar Deposit rate. A comparison with Figure 2 shows the magnitude of the performance drag, which is most severe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

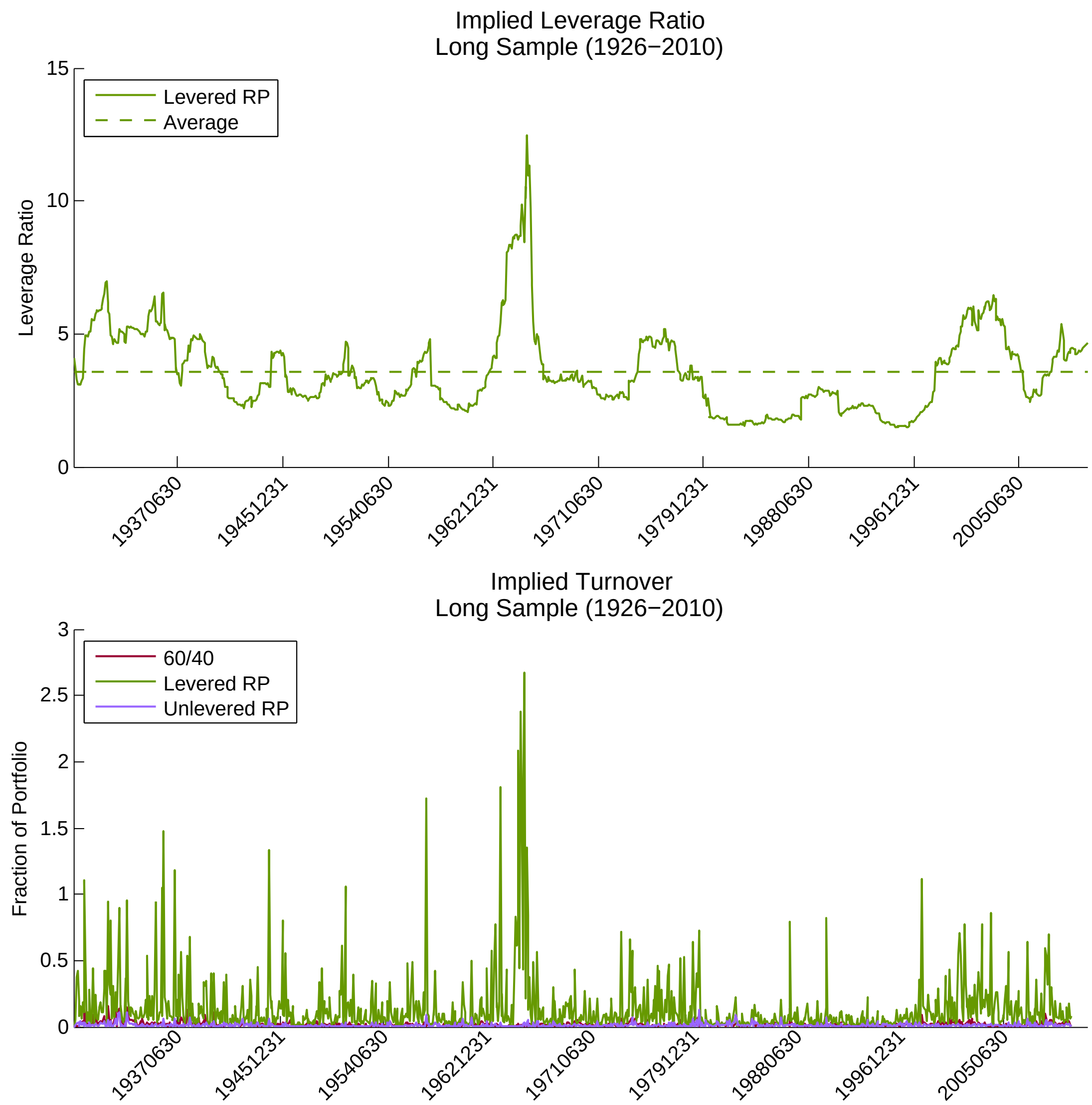

Figure 5: Strategy turnover. The top panel plots the leverage required in order for the estimated volatility of the risk parity strategy to match the estimated volatility of the market at each rebalancing. The average over the entire period is 3.55. The spike in leverage in the mid-1960s corresponds to a rare moment when bond volatility was relative low, and both equity volatility and market weight were relatively high. The bottom panel shows the turnover of the risk parity and 60/40 strategies at each rebalancing.

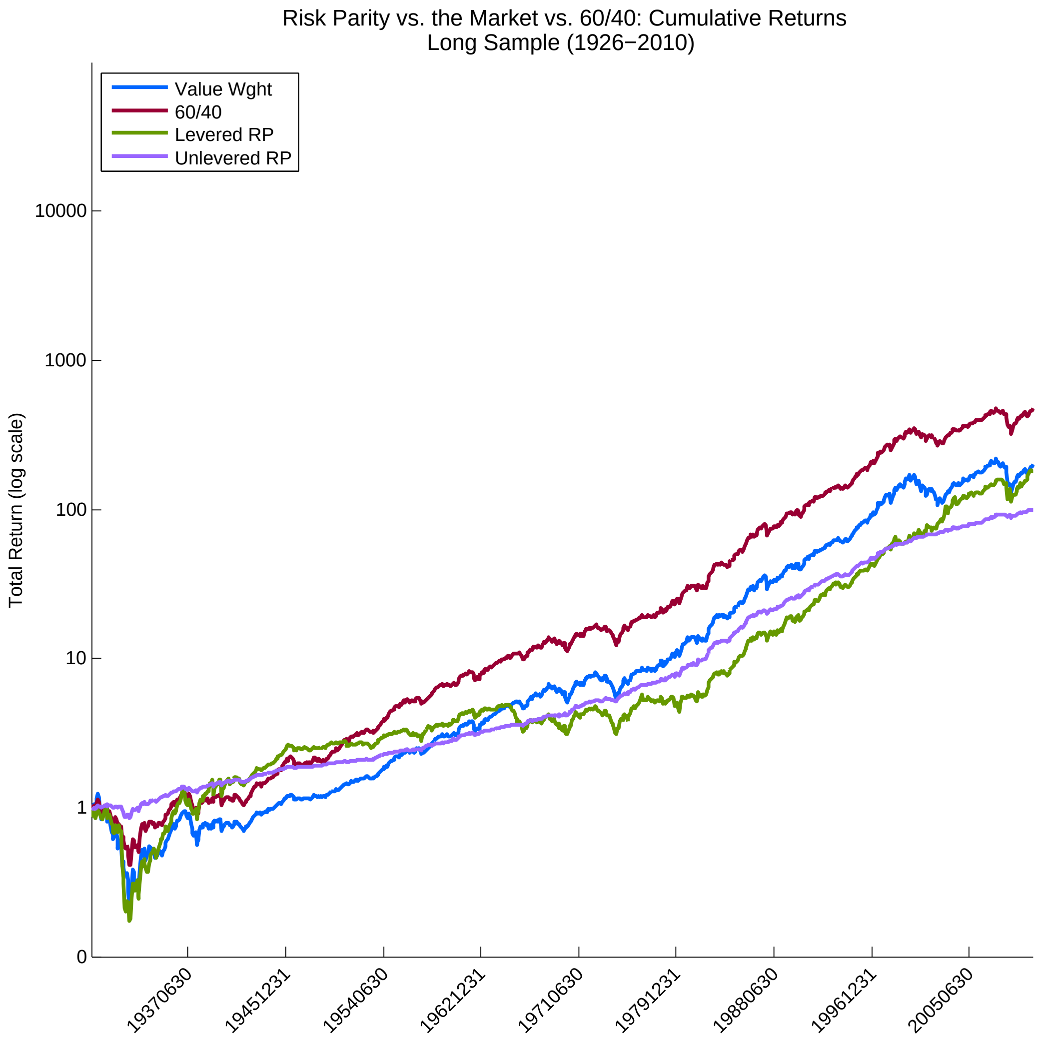

Figure 6: Continuously compounded return to four strategies based on US Equity and US Treasury Bonds over the period 1926–2010. The levered risk parity strategy is financed at the a 3-Month Euro-Dollar Deposit rate and adjustments are made for turnover. A comparison with Figure 3 shows the magnitude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

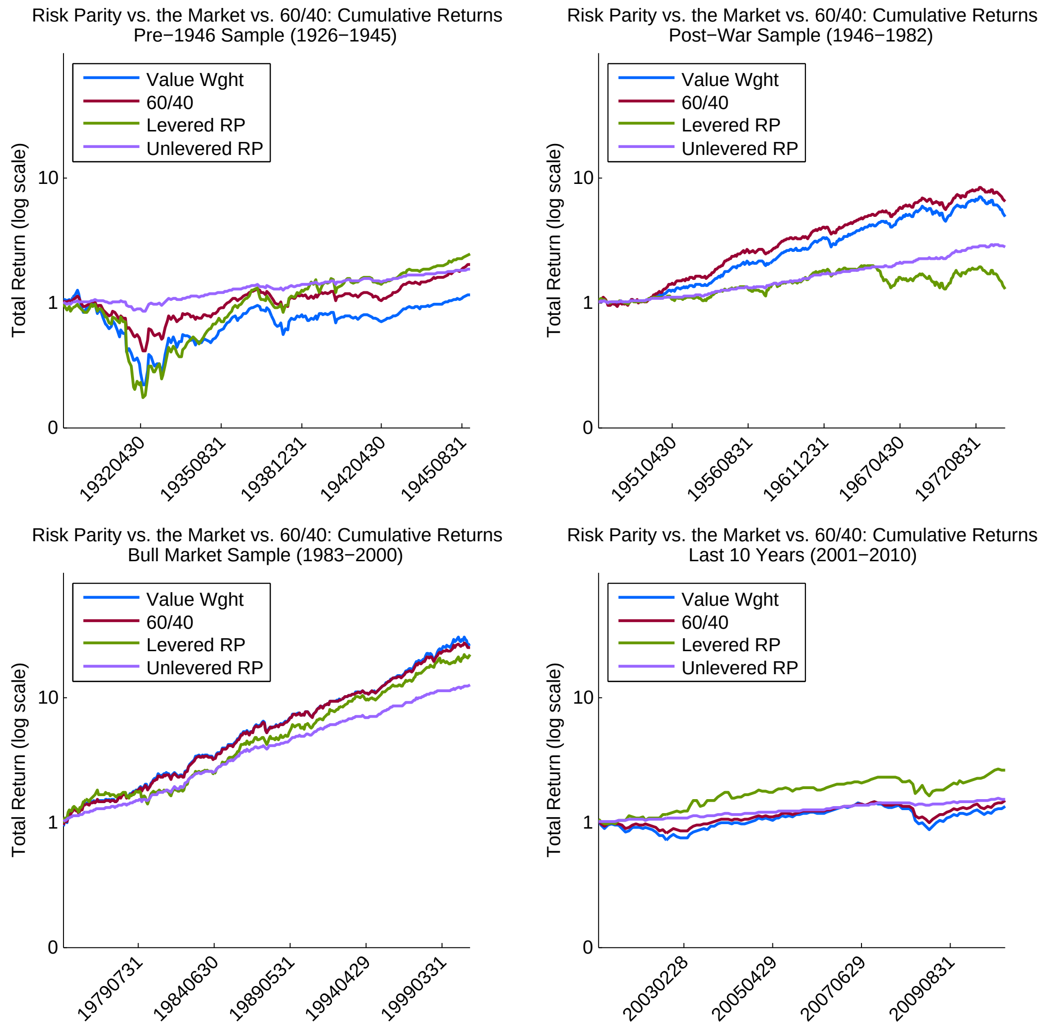

Figure 7: Continuously compounded return to four strategies based on US Equity and US Treasury Bonds over 4 subperiods. The levered risk parity strategy is financed at the a 3-Month Euro-Dollar Deposit rate and adjustments are made for turnover. A comparison with Figure 4 shows the magnitude of the performance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

again. This time, the value weighted and levered risk parity strategies finish in a tie. Figure 7 shows the breakdown into subperiods.

4 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 for these comparisons. Disregarding trading costs and assuming borrowing is at the risk-free rate, the return of levered risk parity exceeds 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 is 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 return of levered risk parity exceeds that of 60/40 by only 29 basis points, and is nowhere close to being statistically significant ($P = 0.40$). The alphas are essentially tied. If we also take into account trading costs, 60/40 beats 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.

5 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.[5]

[5]The Sharpe ratios of the levered and unlevered risk parity strategies do not agree, even when borrowing is at the risk-free rate. This is because the leverage is dynamic. The leverage ratio is chosen at

In the Capital Asset Pricing Model (CAPM), the value weighted portfolio uniquely maximizes the Sharpe ratio over the feasible set of portfolios with holdings limited to the risky assets. The Sharpe ratio is the slope of the *capital market line*, which represents all portfolios that are combinations of the value weighted portfolio and cash. These portfolios have a common Sharpe ratio, but they have different excess expected returns that compensate in a linear way for different levels of volatility.

Now, suppose the rate of borrowing is higher than the risk-free rate. Then the capital market line is kinked. Unlevered portfolios lie on the segment between the risk-free asset and the value weighted portfolio, and they have the same Sharpe ratio as the value weighted portfolio. Levered portfolios lie on a different segment, and they have Sharpe ratios that are *lower* than the Sharpe ratio of the value weighted portfolio. The situation is sketched in Figure 10.

Our findings over the Long Sample, 1926–2010, suggest that the CAPM may not hold, because it appears that the Sharpe ratio of the unlevered risk parity strategy exceeds that of the value weighted portfolio. So now alter the CAPM picture, replacing the value weighted portfolio with the unlevered risk parity portfolio. The "risk parity capital market line" also has a kink that distinguishes levered and unlevered portfolios. Sharpe ratios of levered risk parity portfolios are lower than Sharpe ratios of unlevered risk parity portfolios.

The results in Figures 8 and 9 are consistent with (but do not prove) the hypothesis that risk parity strategies do relatively well in turbulent periods. However, those who measure performance with Sharpe ratio may prefer unlevered risk parity, while those who measure performance with cumulative returns may prefer levered risk parity. Other conclusions may be drawn from Table 1, which displays standard statistics on the four strategies. Investors concerned about skew and kurtosis may find interesting information in the last two columns of the table.

6 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.

---

each monthly rebalancing so that the conditional ex post volatilities of the levered risk parity and value weighted strategies match.

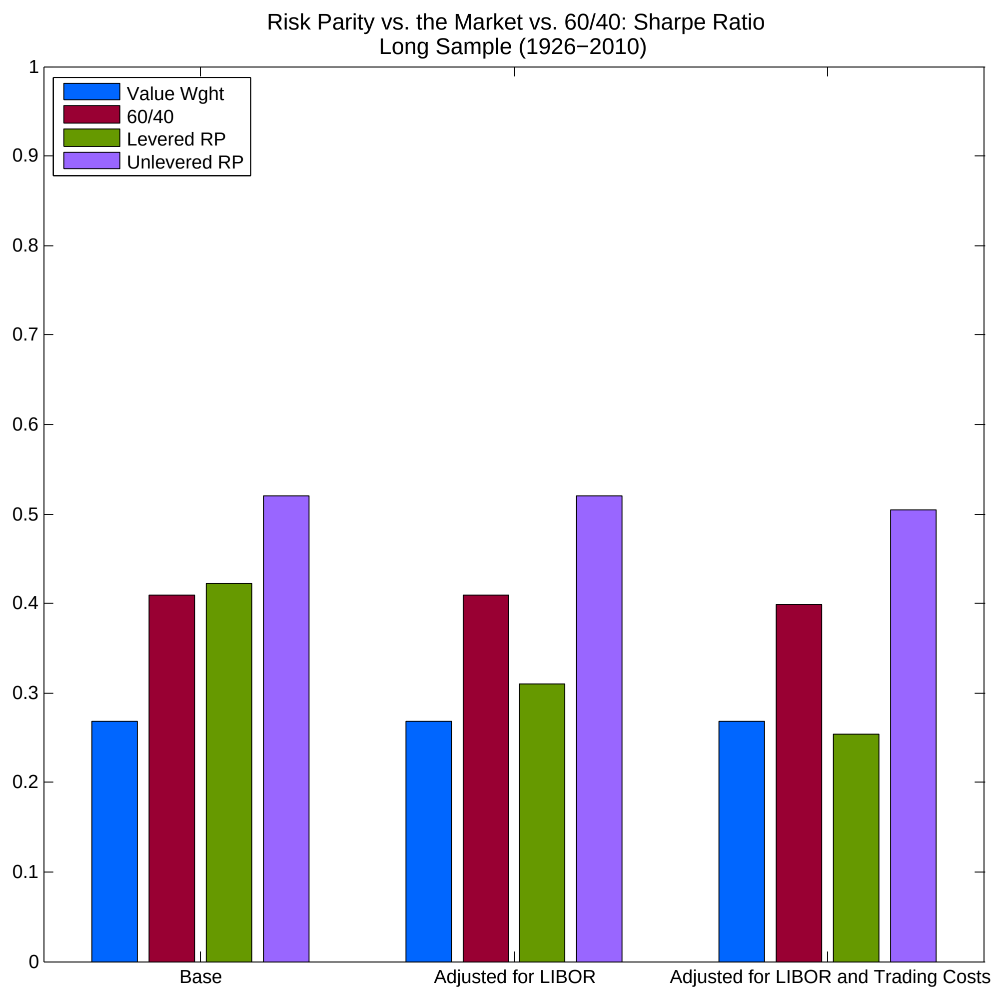

Figure 8: Realized Sharpe ratios for the four strategies over the period 1926–2010. Unlevered risk parity dominates, even before adjustment for market frictions.

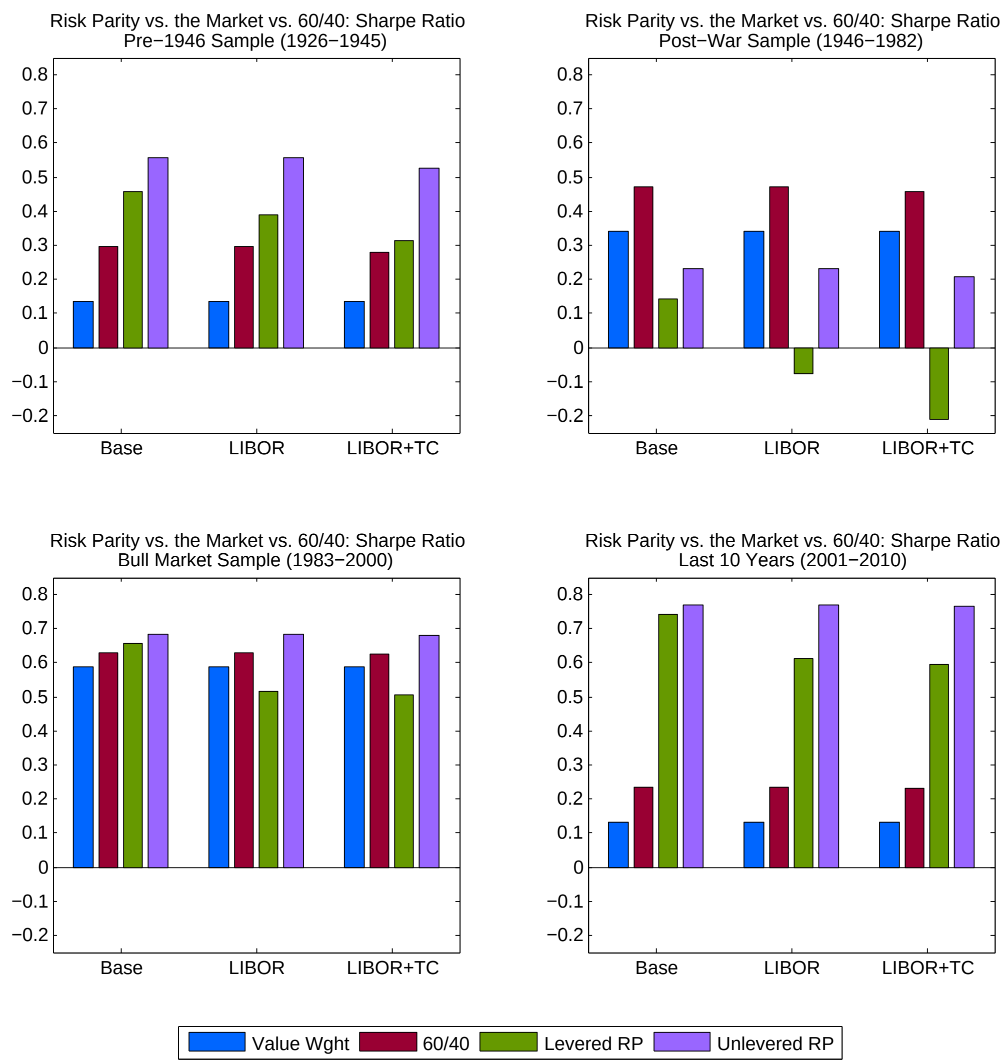

Figure 9: Realized Sharpe ratios for the four strategies over the four subperiods. Apart from the Post-War Sample, Unlevered Risk Parity dominates, even before adjustment for market frictions.

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 caveats. First, levered risk parity underperformed during a relatively long subperiod: the 45-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.

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 sum of the face value outstanding for each bond 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

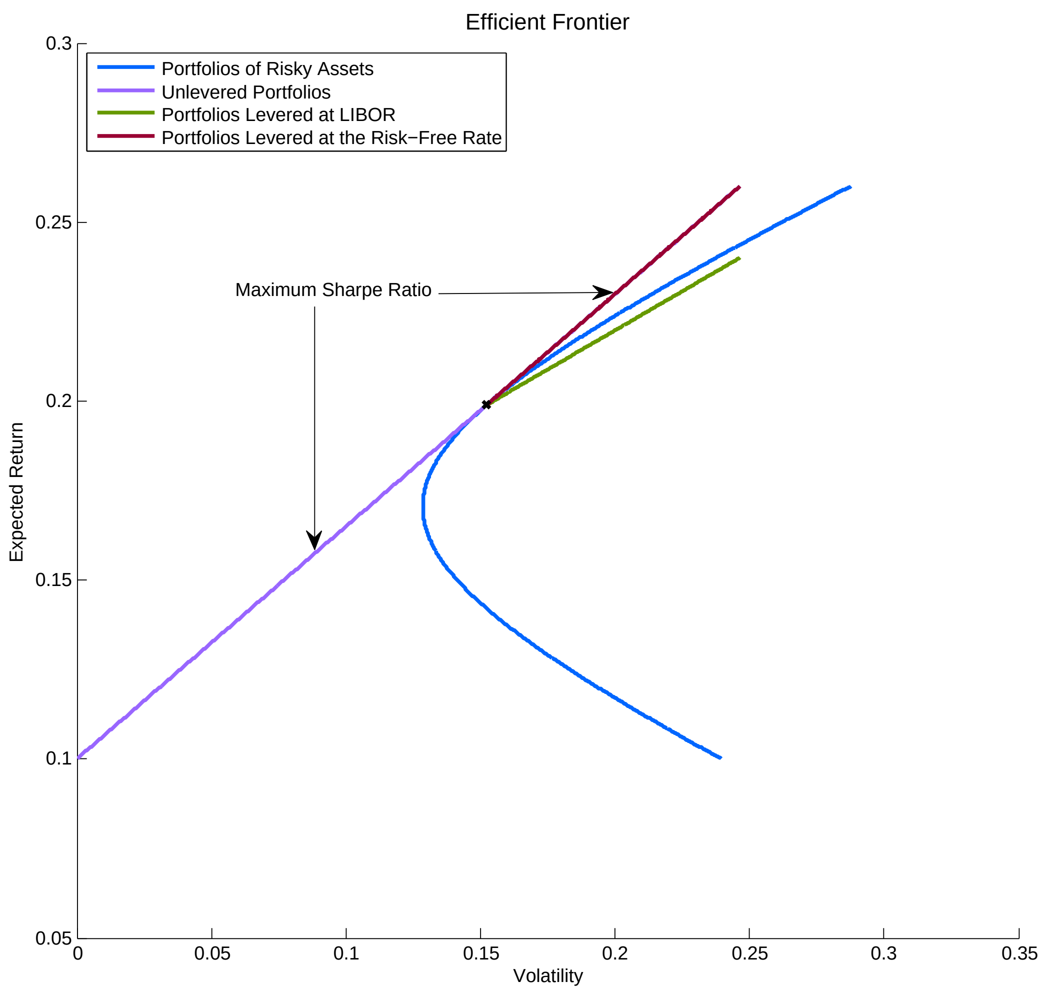

Figure 10: When the rate of borrowing is higher than the risk-free rate, the Capital Market Line in the standard mean-variance diagram is kinked. The ex ante Sharpe ratio of a levered portfolio consisting of the market and cash is lower than the ex ante Sharpe ratio of the market portfolio.

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 LIBOR	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 LIBOR 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.

2010. Prior to January of 1971, a constant of 60 basis points is added to the 90-day T-Bill rate.[6]

B Strategies

To construct an *unlevered* Risk Parity portfolio, at the end of every calendar month the volatilities of each asset class (i.e., stocks and bonds) are estimated using a 36-month rolling window of trailing total returns. Hence, the volatility estimate for asset class i at time t is given by

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

The portfolio weight for each asset class i is then given by

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

where

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

The levered risk parity strategy in this article is conditional, in the sense that its ex post volatility matches the ex post volatility of the market portfolio at each rebalancing. To calculate conditionally levered Risk Parity portfolio weights, the returns to the unlevered Risk Parity portfolio for the trailing 36 months are estimated as

$$r_{u,t-s} = \sum_i w_{i,t} r_{i,t-s},$$

for $s = 1, \ldots, 36$. The leverage ratio required to match the trailing 36-month realized volatility of the value weighted index is estimated as

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

where

$$\begin{aligned} \hat{\sigma}_{v,t} &= \text{std}(r_{v,t-36}, \ldots, r_{v,t-1}), \\ \hat{\sigma}_{u,t} &= \text{std}(r_{u,t-36}, \ldots, r_{u,t-1}), \end{aligned}$$

and $r_{v,t}$ is the return of the value weighted index at time t. The Risk Parity portfolio weights are then given by

$$w_{i,t}^{*} = l_t w_{i,t}.$$

[6] 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.

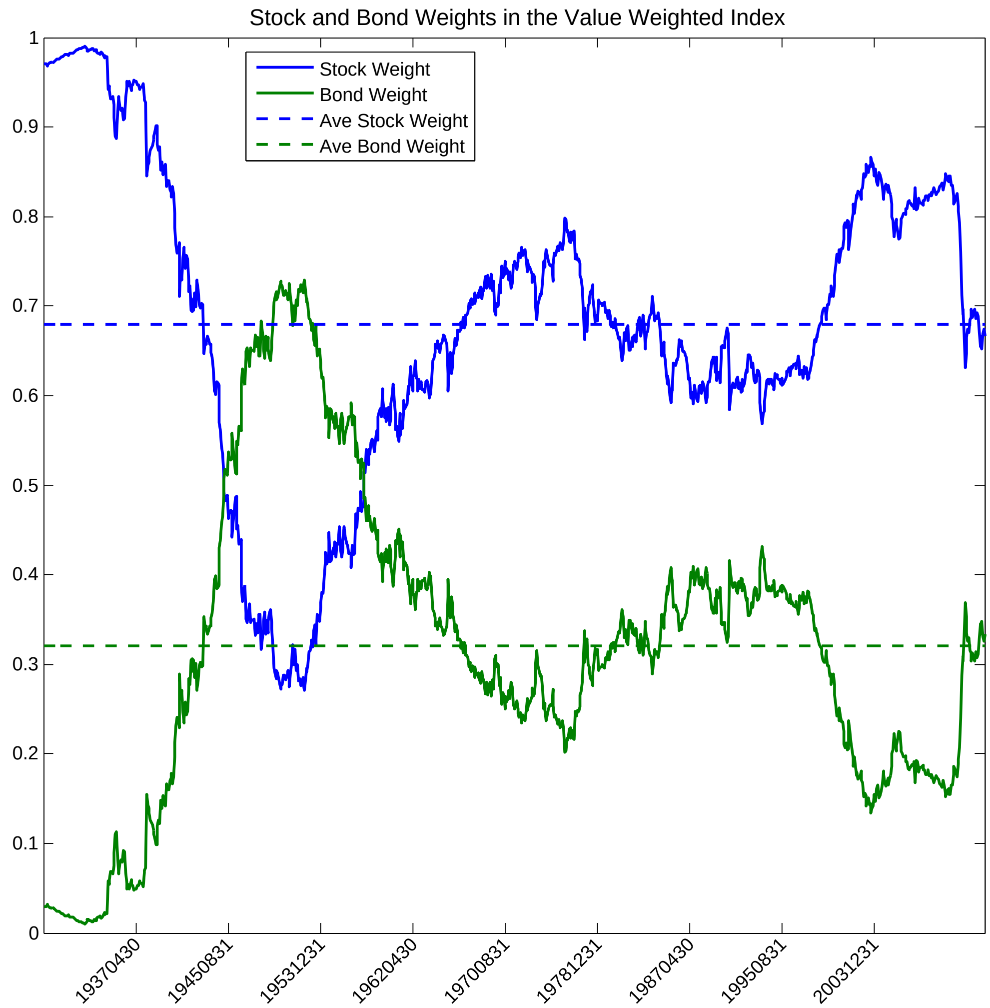

Figure 11: Weights for stocks and bonds implied by market capitalization over the sample period.

The *return* of the conditionally levered Risk Parity portfolio at time t is estimated as

$$\begin{aligned} r_{l,t} &= \sum_i w_{i,t} r_{i,t} + \sum_i (l_t - 1) w_{i,t} (r_{i,t} - b_t) \\ &= \sum_i w_{i,t} r_{i,t} + \sum_i (w^*_{i,t} - w_{i,t})(r_{i,t} - b_t), \end{aligned}$$

where b_t is the borrowing rate at time t. The *excess return* of the conditionally levered Risk Parity portfolio at time t is then given by

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

where $r_{f,t}$ is the risk-free rate at time t.

Asness et al. (2011) implement an unconditional levered risk parity strategy. To calculate unconditionally levered Risk Parity portfolio weights, a constant k is chosen such that

$$w^{**}_{i,t} = k\hat{\sigma}^{-1}_{i,t},$$

$$r^e_{l,t} = \sum_i w_{i,t} r_{i,t} + \sum_i (w^{**}_{i,t} - w_{i,t})(r_{i,t} - b_t) - r_{f,t},$$

and

$$\text{std}(r^e_{l,37}, \ldots, r^e_{l,T}) = \sigma,$$

where σ is a desired target volatility and T is the *last* month in the sample period (i.e., if the sample period is January 1926 through December 2010, then $T = 1020$). One *possible* choice of σ is the realized volatility of the excess return of the value weighted index.

C Trading Costs

In order to estimate trading costs, the *turnover* of a portfolio is calculated as outlined below.

60/40 portfolio:

$$\begin{aligned} \tilde{w}_{s,t} &= \frac{(1 + r_{s,t}) w_{s,t-1}}{(1 + r_{s,t}) w_{s,t-1} + (1 + r_{b,t}) w_{b,t-1}} \\ \tilde{w}_{b,t} &= \frac{(1 + r_{b,t}) w_{b,t-1}}{(1 + r_{s,t}) w_{s,t-1} + (1 + r_{b,t}) w_{b,t-1}} \\ x_t &= |\tilde{w}_{s,t} - w_{s,t}| + |\tilde{w}_{b,t} - w_{b,t}|, \end{aligned}$$

where $w_{s,t}$ ($w_{b,t}$) is the policy weight on stocks (bonds) in period t, $\tilde{w}_{s,t}$ ($\tilde{w}_{b,t}$) is the weight implied by the stock (bond) return from $t-1$ to t, $r_{s,t}$ ($r_{b,t}$) is the stock (bond) return from $t-1$ to t, and x_t is the turnover required to achieve the new policy weights.

Risk Parity portfolio:

$$\begin{aligned}
\tilde{w}_{s,t} &= \frac{(1+r_{s,t})w_{s,t-1}}{(1+r_{s,t})w_{s,t-1}+(1+r_{b,t})w_{b,t-1}} \\
\tilde{w}_{b,t} &= \frac{(1+r_{b,t})w_{b,t-1}}{(1+r_{s,t})w_{s,t-1}+(1+r_{b,t})w_{b,t-1}} \\
x_t &= |w_{s,t}\ell_t - \tilde{w}_{s,t}\ell_{t-1}| + |w_{b,t}\ell_t - \tilde{w}_{b,t}\ell_{t-1}|,
\end{aligned}$$

where $w_{s,t}$ ($w_{b,t}$) is the (unlevered) risk parity weight on stocks (bonds) in period t, $\tilde{w}_{s,t}$ ($\tilde{w}_{b,t}$) is the (unlevered) risk parity weight implied by the stock (bond) return from $t-1$ to t, $r_{s,t}$ ($r_{b,t}$) is the stock (bond) return from $t-1$ to t, ℓ_t is the leverage ratio at time t, and x_t is the turnover required to achieve new (levered) risk parity weights at time 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

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

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 slightly 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 4 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 A Comment on Methodology

The statistics on the value weighted, 60/40 and unlevered risk parity strategies in Panel A of Table 1 are nearly identical to the statistics for these strategies in Panel A of Table 2 of Asness et al. (2011). This is because the data, time horizons, and methodologies are the same. Notably, the statistics on the levered risk parity strategies do not agree. The discrepancy is due to an essential difference in the way the two levered risk parity strategies are executed. To explain the difference between the strategies, we first note their key point of similarity. At each monthly rebalancing, the two strategies assign the same *raw* weight to each asset class. The raw weight is the inverse to the ex post volatility. [7] This raw-weighted portfolio has equal risk contributions from US Equity and US Treasury Bonds. However, the leverage ratio has not yet been set. In the conditional strategy used in this article, the leverage ratio at each monthly rebalancing is chosen so that the conditional ex post volatility estimates of the levered risk parity and value weighted strategies match. Asness et al. (2011) scale the raw weights by a time independent constant chosen so that over the period 1926–2010, the unconditional ex post volatilities of the levered risk parity and value weighted strategies match. This unconditional volatility was not known in 1926. Asness et al. (2011) point out that the Sharpe ratio (and other important statistics) do not depend on the volatility, so long as the levered position is financed at the risk-free

[7] The ex post volatility is estimated from monthly data with a three-year rolling window.

rate.[8] The Sharpe ratio of the levered risk parity strategy in Asness et al. (2011) is .5 over the period 1926–2010 and ours is .4. However, the strategy in Asness et al. (2011) requires that the investor must maintain a constant scaling of the raw weights throughout the period 1926–2010, no matter what the market does. Finally, we note that excess return to the levered risk parity strategy in Asness et al. (2011) depends linearly on the scale factor. Perhaps a fourth essential point to remember when backtesting strategies is that the parameters of a strategy must be determined by past and contemporaneous observables.

References

Clifford Asness, Andrea Frazzini and Lasse Heje Pedersen. Leverage Aversion and Risk Parity, 2011. Working paper, Yale University.

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

Jonah Lehrer. The Truth Wears Off: Is There Something Wrong with the Scientific Method. *The New Yorker*, 2010.

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

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

[8] As noted above, the Sharpe ratio *does* depend on the volatility if leverage is financed at a higher rate.