3

Portfolio Construction and Performance Measurement when Returns are Non-Normal

by

Karen Benson †
Philip Gray †
Egon Kalotay §
Judy Qiu †

Abstract:

The foundation of popular approaches to portfolio construction and performance measurement lies in the mean-variance framework of Markowitz (1952, 1959). However, the suitability of such approaches in practice is questionable in light of considerable evidence of non-normalities in returns. This paper explores the potential usefulness of a non-parametric approach to portfolio construction and performance measurement recently proposed by Stutzer (2000). The Portfolio Performance Index (PPI) is based on the notion that investors associate risk with the failure to achieve a target return. Stutzer proposes that portfolio construction and performance measurement be approached by calculating the decay rate in the probability that a given portfolio will underperform its designated benchmark. By comparing the PPI and Sharpe ratio metrics, this paper presents preliminary evidence of the economic significance of non-normalities in Australian equity returns, and documents the impact of such on portfolio construction and performance evaluation practice.

Keywords:

PERFORMANCE MEASUREMENT; PORTFOLIO CONSTRUCTION; DOWNSIDE RISK; SHARPE RATIO; INFORMATION RATIO.

† UQ Business School, The University of Queensland, St Lucia, QLD 4072. Email: p.gray@business.uq.edu.au
§ Department of Accounting and Finance, Macquarie University.

The authors are grateful to an anonymous referee for constructive suggestions and Mike Stutzer for guidance on computational issues. This project is funded by the Australian Research Council Discovery Project DP0773662.

Australian Journal of Management, Vol. 32, No. 3, Special Issue, March 2008, © The University of New South Wales

1. Introduction

Academics and practitioners alike are keenly interested in measuring the performance of investment portfolios. Over the last five decades, many performance metrics have been proposed and a sizeable literature has developed assessing the ability of managed fund portfolios to outperform benchmarks. The models adopted in academic papers vary in their degree of sophistication, but predominantly focus on some form of risk-adjusted performance.[1] In the financial press, portfolio performance is generally reported as the return in excess of a benchmark. For example, information ratios or Sharpe ratios express a portfolio's excess return relative to its variability.[2] These measures have been enthusiastically embraced by funds management professionals (Goodwin 1998).

More recently, Stutzer (2000) has proposed a new approach to portfolio construction and performance measurement that has both academic rigour and practical appeal. In light of the widespread practice of benchmarking, both investors and fund managers are acutely concerned with fund performance relative to the designated benchmark. In some respects, common usage of the term 'risk' has evolved to represent the probability that an investment will underperform a given target. As a performance metric, Stutzer (2000) derives a portfolio performance index (PPI) that captures this notion of risk. The intuition behind the PPI is relatively straight forward and easily interpreted by the layperson. This feature is particularly important. Recent legislative changes regarding retirement savings have prompted both a rapid expansion of the managed fund industry and a heightened level of interest by unsophisticated investors.

The purpose of this paper is to explore the potential usefulness of Stutzer's PPI methodology via a number of applications to the Australian market. We first demonstrate the application of the PPI to optimal portfolio construction. Similar to the way in which portfolios can be constructed to maximise the Sharpe ratio, stock weights can be selected to maximise the PPI. The empirical example illustrates this process and compares optimal PPI weights to a number of common alternatives.

Our second application uses the PPI to rank the performance of a sample of Australian equity trusts. The extent to which rankings under Stutzer's PPI differ from rankings under the Sharpe ratio is an indication of the economic significance of return deviations from normality. In the presence of departures from normality, and assuming that the PPI metric is more-closely aligned with investors' objective function, our preliminary results suggest that a focus on the Sharpe ratio in the Australian context may be misleading.

The remainder of the paper is organised as follows. Section 2 briefly reviews the mean-variance approach to performance measurement and portfolio construction and, in doing so, motivates the development of and intuition for the PPI. Section 3 introduces Stutzer's (2000) portfolio performance index, providing an overview of key formulae. Section 4 describes the data and methodology

1. See, for example, Jensen (1968), Treynor and Mazuy (1966), Henriksson and Merton (1981), Carhart (1997), Daniel, Grinblatt, Titman and Wermers (1997), and Ferson and Schadt (1996) for international findings and Sawicki and Ong (2000) for Australian findings.
2. Popular practice uses the term Sharpe ratio when a riskfree benchmark is adopted and Information ratio when an index benchmark is chosen (see Goodwin 1998). This paper uses the term Sharpe ratio to represent both choice of benchmarks.

employed in the study. Section 5 presents the empirical analysis and section 6 concludes the paper.

2. Performance Measurement and Portfolio Construction

The Sharpe (1966) statistic measures the performance of a fund calculated as the ratio of the reward from holding a portfolio to the risk of the portfolio—in essence, the reward (in excess of some benchmark) per unit of risk. Originally named the reward-to-variability ratio, the measure has almost universally become known as the Sharpe ratio.

Ex post performance measurement is often conducted by ranking portfolios in descending order of their Sharpe ratio. Similarly, portfolio construction can be approached by selecting the combination of assets that maximises the ex ante Sharpe ratio. Under certain assumptions (discussed shortly), the Sharpe-ratio-maximising portfolio is a mean-variance efficient portfolio. Formulae giving the requisite weights to be invested in each asset are easily derived and provided in most elementary investment texts (e.g. Elton & Gruber 1995; Bodie, Kane & Marcus 2005).

The origins of the Sharpe ratio are deeply rooted in the mean-variance paradigm of Markowitz (1952, 1959), whereby investors are assumed to care only about the expected return and variance of asset returns. Investors can limit their focus to the first two moments of the return distribution either when either returns are normally distributed, or when they have a quadratic utility function.[3] Since the notion of quadratic utility is unappealing, the mean-variance framework is generally predicated on the assumption that returns are normally distributed.[4]

Despite widespread usage of the Sharpe ratio, there is an overwhelming body of literature suggesting that stock returns are not normally distributed. Consistent with the findings of international studies, the Australian evidence against normality dates back at least three decades (Praetz & Wilson 1978; Beedles 1986; Beedles, Dodd & Officer 1988; Alles & Spowart 1995; Gray, Kalotay & McIvor 1998). Further, even if stock returns are normally distributed, the use of derivatives is likely to induce non-normality into the distribution of portfolio returns. In the absence of normally distributed returns, investors will be concerned with higher moments of the return distribution (e.g. skewness). The common use of protective put strategies suggests a clear preference for positive skewness.

The practical usefulness of the Sharpe ratio might also be questioned on the basis of the appropriate definition of 'risk'. The Markowitz (1952, 1959) framework uses the variance (or standard deviation) of asset returns as the risk measure, thereby placing equal emphasis on return deviations below *and above* the mean. As such, a risk-averse investor in Markowitz's mean-variance world views downside and upside variation with equal distaste (Leland 1999).

3. In fact, normally-distributed returns is a sufficient but not necessary assumption. The mean-variance framework obtains so long as investors are risk averse and returns are generated by a spherically-invariant class. (e.g. Cauchy).
4. Quadratic utility functions are problematic from a number of perspectives. They involve several parameters that are difficult/impossible to estimate. Most importantly, quadratic utility implies that the investor reaches a satiation point, beyond which increased wealth/consumption reduces utility.

Despite the focus on the mean-variance framework, alternate measures of risk have long been recognised. Markowitz (1959) first suggested the use of semi-variance to measure the degree of variation below the mean return (downside risk). Sortino and van der Meer (1991) propose a modification to the Sharpe ratio whereby the (lower) semi-variance replaces standard deviation in the denominator. The recent advent of the Value-at-Risk measure is a further indication of the asymmetric nature of investor concern over risk. Ang, Chen and Xing (2006) argue that investors demand additional compensation to hold stocks with higher sensitivities to downside market movements and document a downside risk premium in the order of 6% per annum.

The PPI measure proposed by Stutzer (2000) is based on the notion that investors tend to associate risk with the failure to achieve a target return. The widespread practice of benchmarking a portfolio to a specified index lends credence to this interpretation of risk. Stutzer (2000) shows that, when a portfolio is expected to earn a higher average return than the chosen benchmark, the probability that the portfolio will underperform the benchmark approaches zero as the sample period lengthens. Further, the rate at which the probability of underperformance declines is a computable exponential decay rate. Stutzer (2000) proposes that this decay rate be adopted as the measure of portfolio performance and shows how to construct a portfolio to maximise the decay rate.

Stutzer's portfolio performance index has a number of desirable features. First, unlike the Markowitz mean-variance paradigm that gives rise to the Sharpe ratio, it is a distribution-free measure of performance. The probabilities inherent in the index calculation rely only on the central limit theorem applied to portfolio returns. Second, the portfolio performance index has a simple and intuitive interpretation that is likely to appeal to both fund managers and investors. Investors often view the return on the chosen benchmark as the minimum acceptable performance for their fund—after all, the investor is assured of zero underperformance if she invests in the benchmark index. As such, an investor expecting fund-manager skill to generate outperformance will take a dim view of benchmark underperformance. Similarly, fund managers are aware of the adverse implications of underperformance on fund flows and most likely their remuneration. Since the portfolio performance index explicitly quantifies the decay rate in the probability that the time-averaged portfolio return will underperform the benchmark, this metric has considerable practical appeal to investment market participants (both sophisticated and unsophisticated).

Third, as Stutzer (2000) explains, the portfolio performance index captures the investor preference for positive skewness that the mean-variance framework disregards. At an elementary level, stocks with positively skewed returns may be more attractive under the PPI during portfolio construction than the mean-variance framework recognises. Ultimately, however, it is the joint distribution of stock returns and resulting skewness in portfolio returns that influences the likelihood of benchmark underperformance. The PPI approach implicitly accommodates these factors. Finally, the portfolio performance index retains the ease of implementation of the Sharpe ratio, with arguably a more-intuitive interpretation. While the formulae in the next section show that the PPI metric requires solving an optimisation problem, this is a relatively simple task which can be conducted in Microsoft Excel.

3. The Portfolio Performance Index

Assume that the portfolio of interest is benchmarked against another asset (for example, a market index or the riskfree rate). Let $R_{p,t}$ denote the time t return on portfolio p in excess of the chosen benchmark. The average excess return to portfolio p over T sample periods is simply:

$$\overline{R}_p(T) = \frac{1}{T}\sum_{t=1}^{T} R_{p,t}\,. \tag{1}$$

Stutzer (2000) highlights that, if the portfolio has a positive expected excess return, the law of large numbers implies that the probability of observing a negative sample excess return $\left(\overline{R}_p(T)\right)$ approaches zero as the sample period T increases. Further, this probability converges to zero at an exponential rate (I_p):

$$\Pr\left(\overline{R}_p(T) \le 0\right) \approx \frac{c}{\sqrt{T}} e^{-I_p T}\,. \tag{2}$$

In light of the preceding discussion of aversion to downside (or underperformance), Stutzer (2000) argues that this limiting relationship may find useful application in portfolio construction and performance measurement. For example, a fund manager who is averse to underperformance might construct a portfolio such that the probability of earning a non-positive average excess return goes to zero as quickly as possible. Similarly, investors may rank fund performance in decreasing order of I_p.

The exponential decay rate (I_p), which Stutzer (2000) refers to as the 'portfolio performance index', is given by:

$$I_p = \max_{\theta}\left(-\ln E\left(e^{\theta \overline{R}_p(T)}\right)\right), \tag{3}$$

where $\theta < 0$. When calculated ex post, I_p provides an alternative performance metric to the Sharpe ratio. I_p estimates the rate at which the probability of underperformance decays to zero. When investors are averse to benchmark underperformance, the portfolio with the fastest decay rate is preferred. Stutzer shows that, when stock returns are in fact normally distributed, the portfolio performance index (I_p) is a one-to-one function of the traditional Sharpe ratio; hence, portfolio rankings will be identical under either metric. In the absence of normality, the ranking of portfolios by I_p reflects a preference for positive skewness in portfolio returns.

Similarly, the portfolio performance index can be utilised in the ex ante construction of portfolios. Assume that there are N assets under consideration for inclusion in the portfolio, and that a time series of T excess return observations ($R_{i,t}$) exists for each asset i. At any time t, the portfolio excess return is calculated:

$$R_{p,t} = \sum_{i=1}^{N} w_i R_{i,t}\,, \tag{4}$$

where w_i denotes the portfolio weighting assigned to asset i. The sample estimate of the right-hand side of equation (3) is calculated by:

$$\widehat{I}_p = \max_{\theta}\left(-\ln \frac{1}{T}\sum_{t=1}^{T} e^{\theta R_{p,t}} \right). \tag{5}$$

Finally, the optimal stock weightings under the portfolio performance index metric are estimated using the following maximisation problem:

$$I_m = \max_{w_1,\ldots,w_n} \max_{\theta}\left(-\ln \frac{1}{T}\sum_{t=1}^{T} e^{\theta R_{p,t}} \right). \tag{6}$$

Warning that good initial guesses are important, Stutzer (2000) suggests using the Sharpe-ratio-maximising weights ($w_1,\ldots, w_N$) as an initial guess in equation (6). Similarly, a reasonable starting value for θ is the negative of the mean excess return divided by its variance.

4. Data and Methodology

The empirical study is divided into two components. We begin by examining the optimal portfolio weights assigned to stocks under alternate portfolio-formation approaches. First, the optimal portfolio under the Sharpe ratio is identified using the formula from mean-variance theory (e.g. Stutzer 2000). Second, the portfolio given by equation (6) is identified using numerical optimisation. This portfolio is optimal in the sense that it assigns weights to stocks such that the probability of the portfolio underperforming the chosen benchmark decays to zero as quickly as possible.

A third relevant alternative is to select portfolio weights to maximise the Sortino ratio, which modifies the Sharpe ratio by using the (square root of the) lower semi-variance in the denominator. The lower semi-variance of returns is calculated in a similar fashion to variance, with the exception that only returns falling below the mean are included. By using the downside risk, the Sortino ratio reflects investors' concern over downside and induces a preference for positively skewed stocks. The Sortino-ratio-maximising portfolio weights are obtained using numerical optimisation.

Data for the first component of the analysis are drawn from the Australian Graduate School of Management's Centre for Research in Finance (CRIF) database which contains monthly data for all stocks listed on the Australian Securities Exchange (ASX). The top 200 stocks by market capitalisation as at December 2005 are identified. From this list, twenty stocks are chosen at random for the empirical analysis.[5] For each stock, a timeseries of 120 monthly returns from 1996–2005 is obtained.

Two alternate benchmarks are studied. First, optimal portfolio construction is considered with the portfolio benchmarked to the riskfree rate. The proxy for the

5. While fund managers are likely to hold more than twenty stocks in a portfolio, we limit the analysis to twenty stocks simply for convenience in tabulating the results.

riskfree rate is the return on 13-week Treasury notes supplied by CRIF. Second, the portfolio is benchmarked to the value-weighted market index supplied by CRIF.

The second component of the empirical study examines the ranking of fund performance under alternate metrics. Fund return data are sourced from Morningstar, accessed through the Morningstar Direct internet-based platform. The analysis focuses on the set of unit trusts whose products are categorised as Australian equity investments. This sample includes retail and wholesale funds, but excludes specialist funds (for example, natural resources and real estate). The resulting sample is a homogenous group of funds for which benchmarking to the ASX200 market index is reasonable.

Over the period January 2002 to December 2006, monthly returns are available for 295 unit trusts represented by 60 fund management firms.[6] For each fund, Stutzer's portfolio performance index (5) is calculated using numerical optimisation. In addition, the Sharpe ratio (relative to the ASX200 market index) is calculated for each fund.

5. Results and Discussion

5.1 Portfolio Construction Under Asymmetric Returns

Table 1 presents the twenty stocks considered in the study, along with the CRIF group tcode. Summary statistics are calculated using monthly continuously-compounded returns (in excess of the riskfree rate) over the period 1996–2005 (120 months). The sample stocks display a mix of mean/standard deviation characteristics. The marginal return distributions display a clear predominance of negative skewness and positive excess kurtosis.[7]

The shaded region in table 1 reports the optimal portfolio weights under each approach to portfolio formation, where weights are unconstrained (i.e. short selling is possible). Portfolio holdings are ordered in descending order of the Sharpe-ratio-maximising weights. Comparing the optimal weights under the Sharpe ratio and Stutzer's Portfolio Performance Index (PPI), table 1 suggests that weights are qualitatively similar with modest variations. Compared to the Sharpe ratio, the PPI underweights Milton Corp and overweights Reece Australia. This is consistent with Stutzer's conjecture that the PPI captures a preference for positive skewness (MLT1 is negatively skewed while REH1 is positively skewed). However, the relationship between the Sharpe ratio and PPI weights is not strictly driven by skewness. Leighton Holdings, for example, is overweighted by the PPI despite its negative skewness. Similarly, Patrick Corp is positively skewed yet underweighted by PPI.

---

6. The Morningstar return calculation is a return percentage taking into account the change in price, reinvesting and distribution during the month. The returns are net of monthly management and administration fees, but not load fees.
7. The excess kurtosis is calculated relative to a Normal distribution which has kurtosis of 3.0 Positive excess kurtosis signals a return distribution with tails fatter than a Normal. QBE Insurance's excess kurtosis of 20.94 stands out. This extreme statistic is solely attributable to a negative monthly return of nearly 60% surrounding the September 11 terrorist attacks. Removing this single observation, QBE's kurtosis is more typical 0.36.

Table 1

Stock Characteristics and Portfolio Weights With Riskfree Benchmark

Twenty stocks are randomly selected from the Top 200 stocks by market capitalisation as at December 2005. Summary statistics are for monthly continuously-compounded returns (in excess of the riskfree rate) over the period 1996–2005 (120 months). The optimal portfolio weights are calculated by three alternate approaches: (i) to maximise the Sharpe ratio; (ii) to maximise the Portfolio Performance Index; and, (iii) to maximise the Sortino ratio. Under each approach, weights are first estimated without short-selling constraints, then constrained to be non-negative. The designated benchmark for the portfolio is the riskfree rate supplied by CRIF proxied by 13-week Treasury notes.

						Unconstrained Optimal Portfolio Weights (%)			Constrained (no short selling) Optimal Portfolio Weights (%)		
Stock	CRIF tcode	Mean	StdDev	Skewness	Excess Kurtosis	Sharpe Ratio	Performance Index	Sortino Ratio	Sharpe Ratio	Performance Index	Sortino Ratio
Milton Corp	mlt1	0.0120	0.0324	–0.19	1.69	46.57	43.97	45.48	37.49	40.58	42.47
Reece Australia	reh1	0.0187	0.0503	0.38	1.03	37.53	39.15	41.06	33.58	33.82	36.67
Commonwealth Bank Australia	cba1	0.0118	0.0520	–0.28	0.48	19.00	24.60	26.95	11.72	11.57	9.84
Leighton Holdings	lei1	0.0133	0.0787	–0.96	1.80	8.71	10.29	12.50	4.23	3.33	2.93
United Group	ucr1	0.0209	0.0972	–0.48	1.91	7.80	8.05	5.87	5.13	4.63	4.98
Investa Property	wpr1	0.0044	0.0367	–0.24	–0.14	7.50	8.61	7.79	1.39	0	0
Hills Industries	hil1	0.0122	0.0579	–0.01	0.71	6.05	7.75	6.87	2.58	2.01	0.61
James Hardie Industries	hah1	0.0104	0.0785	–0.22	0.34	4.69	4.71	7.58	0	0	0
QBE Insurance	qbe1	0.0123	0.0844	–3.02	20.94	4.03	2.64	4.17	0	0	0
Patrick Corp	lac1	0.0184	0.0872	0.69	1.87	3.78	2.80	–0.18	3.71	4.06	2.50
Paladin Resources	pdn1	0.0246	0.2692	0.74	2.45	1.20	1.45	1.71	0.17	0	0
Fernz Corp	fzc1	0.0093	0.0808	–0.21	0.14	–0.90	0.93	4.60	0	0	0
CSR Ltd	csr1	0.0101	0.0702	–0.12	0.03	–2.20	–0.01	0.78	0	0	0
Adelaide Brighton Cement	abc1	0.0046	0.0990	0.04	1.11	–3.05	–2.31	–3.96	0	0	0
Seven Network	sev1	0.0055	0.0786	–0.08	0.41	–3.49	–4.66	–7.96	0	0	0
Oilsearch Ltd	osh1	0.0061	0.1241	–0.15	0.39	–4.28	–7.66	–8.36	0	0	0
Western Mining	wmc1	0.0022	0.0830	0.03	1.09	–6.11	–8.65	–8.57	0	0	0
Brambles Industries	bil1	0.0065	0.0766	–1.50	6.64	–6.19	–7.88	–9.45	0	0	0
Brickworks Ltd	bkw1	0.0109	0.0653	–0.33	3.75	–7.41	–8.54	–10.17	0	0	0
Argo Investments	arg1	0.0080	0.0421	–0.08	5.07	–13.24	–15.23	–16.70	0	0	0
						100%	100%	100%	100%	100%	100%

These findings are not unique to the current study. While Stutzer (2000) emphasises the overweighting of positively-skewed stocks, his sample contains a number of examples of positively-skewed stocks being *under*-weighted. As noted earlier, however, it is the joint distribution of stock returns and resulting skewness in portfolio returns that ultimately influences the likelihood of benchmark underperformance on which the PPI metric is based. At the aggregate level, the PPI does indeed construct a portfolio that exhibits positive skewness. Table 2 reports summary statistics for the portfolios dictated by each metric. While the Sharpe-ratio-maximising portfolio has negatively skewed returns (–0.1069), the PPI-maximising portfolio has positive skewed returns (0.1627). The mean-variance-kurtosis characteristics of optimal portfolios under the Sharpe ratio and PPI are quite similar; they differ only in their skewness.

Table 2

Summary Statistics for Optimal Portfolios

Optimal portfolio weights are calculated by three alternate approaches: (i) to maximise the Sharpe ratio; (ii) to maximise the Portfolio Performance Index; and, (iii) to maximise the Sortino ratio. Under each approach, weights are first estimated without short-selling constraints, then constrained to be non-negative. This table reports summary statistics for each of these portfolios. The mean, standard deviation, skewness and excess kurtosis are for the monthly portfolio returns over 1996–2005. Similarly, the Sharpe ratio, Portfolio Performance Index and Sortino ratio are calculated for each portfolio over the same time period.

	Case 1: Riskfree Benchmark Portfolio Construction Approach			Case 2: Market Index Benchmark Portfolio Construction Approach		
	Sharpe	PPI	Sortino	Sharpe	PPI	Sortino
Scenario 1: Portfolio Weights Unconstrained						
Mean	0.0174	0.0180	0.0184	0.0098	0.0102	0.0105
Std Dev	0.0307	0.0322	0.0337	0.0188	0.0199	0.0208
Skewness	–0.1069	0.1627	0.3046	–0.1707	0.1078	0.2261
Excess Kurtosis	2.8126	2.6858	2.7835	3.0988	2.8328	3.0681
Portfolio Sharpe Ratio	0.5662	0.5586	0.5469	0.5199	0.5141	0.5057
Portfolio PPI (Decay Rate)	0.1593	0.1639	0.1598	0.1322	0.1364	0.1340
Portfolio Sortino Ratio	0.5767	0.5955	0.6271	0.5228	0.5343	0.5501
Scenario 2: Portfolio Weights Constrained						
Mean	0.0149	0.0150	0.0151	0.0086	0.0088	0.0089
Std Dev	0.0285	0.0287	0.0290	0.0171	0.0177	0.0183
Skewness	–0.5016	–0.4701	–0.4453	–0.2294	–0.0052	0.1623
Excess Kurtosis	3.4138	3.3582	3.4287	2.8503	2.4506	2.5473
Portfolio Sharpe Ratio	0.5223	0.5217	0.5199	0.5035	0.4990	0.4874
Portfolio PPI (Decay Rate)	0.1267	0.1271	0.1264	0.1236	0.1269	0.1243
Portfolio Sortino Ratio	0.4889	0.494	0.4990	0.4884	0.5016	0.5288

Returning to table 1, the Sortino-ratio-maximising weights are also qualitatively similar to the alternatives. Like the PPI, the Sortino ratio metric is expected to induce positive skewness into the portfolio returns. While, at the individual stock level, there is no clear pattern of overweighting positively-skewed stocks, table 2 reports that the Sortino-ratio-maximising weights do achieve a portfolio with positively-skewed returns (0.3046)—in fact, more so than the PPI (0.1627).

Table 2 also documents the difference in Sharpe ratios, PPIs and Sortino ratios of the three alternate portfolio-construction approaches. By construction, the PPI-maximising portfolio has an underperformance probability that decays to zero faster than the alternatives (0.1639). This decay rate is more than double the equivalent estimate of Stutzer's (2000) riskfree-benchmarked portfolio (0.064). Most likely, this is attributable to the relatively low interest-rate environment over the current sample period—with a low benchmark, the probability of underperformance goes to zero quickly. The Sortino portfolio has a decay rate (0.1598) slightly faster than the Sharpe portfolio (0.1593), which is not surprising given the positive skewness of its returns.

The far-right columns in table 1 present optimal portfolio weights when short-selling is prohibited. While two or three stocks continue to dominate the portfolio, it is noticeable that each method now selects relatively few stocks for inclusion in the portfolio. The short-selling restriction also impacts on the summary statistics for optimal portfolios (table 2). The PPI and Sortino approaches are no longer able to generate portfolios with positively skewed returns. As a consequence, the decay rates and Sortino ratios under the PPI and Sortino approaches drop significantly. This finding demonstrates the importance of short selling to a fund manager concerned with managing downside risk.

Table 3 presents an analysis similar to table 1, with the exception that portfolios are benchmarked to the CRIF value-weighted market index. Examining the unconstrained portfolio weights (shaded region), similar conclusions can be drawn. Relative to Sharpe weights, the PPI over (under) weights some stocks with positive (negative) skewness, yet no consistent pattern is present. Still, the overall returns to the PPI-maximising portfolio are positively skewed (0.1078 in table 2). Naturally, the PPI decay rates of all portfolios are slower when benchmarked to the market index, although not by the magnitude one might have anticipated.[8] For example, the PPI-maximising portfolio has a decay rate of 13.64% per month. Using the rule of 72 approximation, the underperformance probability of this portfolio halves roughly every 5.3 months.[9] The impact of restricting portfolio weights to be positive is less-noticeable when the portfolio is benchmarked to the VW market index (.cf. benchmarked to the riskfree rate).

Comparing tables 1 and 3, one result is striking. When portfolios are benchmarked to the riskfree rate, table 1 portfolio holdings under each approach are highly concentrated in a handful of stocks. Over 80% of the portfolio is held in the two stocks (Milton Corp and Reece Australia). However, when portfolios are benchmarked to the market index, table 3 portfolio holdings are spread across a greater number of stocks (roughly 40% is concentrated in the top two stocks).

8. We note that Stutzer (2000) does not report results for portfolios benchmarked to a market index.
9. See Foster and Stutzer (2004, p. 18) for a discussion of the rule of 72 approximation.

Table 3

Stock Characteristics and Portfolio Weights With Market Benchmark

Twenty stocks are randomly selected from the Top 200 stocks by market capitalisation as at December 2005. Summary statistics are for monthly continuously-compounded returns (in excess of the market index) over the period 1996–2005 (120 months). The optimal portfolio weights are calculated by three alternate approaches: (i) to maximise the Sharpe ratio; (ii) to maximise the Portfolio Performance Index; and, (iii) to maximise the Sortino ratio. Under each approach, weights are first estimated without short-selling constraints, then constrained to be non-negative. The designated benchmark for the portfolio is the CRIF value-weighted market index.

						Unconstrained Optimal Portfolio Weights (%)			Constrained (no short selling) Optimal Portfolio Weights (%)		
Stock	CRIF Tcode	Mean	StdDev	Skewness	Excess Kurtosis	Sharpe Ratio	Performance Index	Sortino Ratio	Sharpe Ratio	Performance Index	Sortino Ratio
Reece Australia	reh1	0.0134	0.0514	0.61	0.33	23.42	27.36	30.60	20.95	23.75	26.61
Milton Corp	mlt1	0.0068	0.0327	–0.38	2.05	20.54	18.68	13.25	15.71	14.49	8.60
Commonwealth Bank Australia	cba1	0.0066	0.0456	–0.12	0.69	14.52	17.66	18.63	12.07	12.85	14.23
QBE Insurance	qbe1	0.0071	0.0725	–3.29	23.42	9.46	7.35	7.40	7.87	6.03	5.88
United Group	ucr1	0.0157	0.0905	–0.51	2.00	9.28	9.54	9.83	7.36	7.36	7.80
Patrick Corp	lac1	0.0132	0.0809	0.68	1.67	9.22	9.26	9.50	8.47	8.46	9.14
CSR Ltd	csr1	0.0049	0.0597	–0.02	–0.17	6.58	5.47	5.16	6.65	6.41	5.56
Seven Network	sev1	0.0003	0.0694	0.46	1.42	5.23	4.94	6.10	4.67	4.28	5.40
Leighton Holdings	lei1	0.0081	0.0739	–0.80	1.46	4.62	5.44	5.05	3.26	3.63	2.52
Western Mining	wmc2	–0.0031	0.0709	0.08	1.09	4.61	5.82	7.61	4.28	5.02	6.77
James Hardie Industries	hah1	0.0051	0.0734	–0.10	0.31	3.95	4.27	3.31	2.63	2.30	2.34
Hills Industries	hil1	0.0070	0.0561	–0.02	0.35	2.37	3.07	5.39	1.33	1.42	3.35
Oilsearch Ltd	osh1	0.0008	0.1114	–0.07	0.45	2.26	1.47	1.38	1.78	1.12	0
Paladin Resources	pdn1	0.0194	0.2618	0.80	2.60	1.86	1.97	1.68	1.80	1.84	1.78
Fernz Corp	fzc1	0.0040	0.0726	–0.09	0.11	1.35	1.30	0.77	1.17	1.03	0
Brambles Industries	bil1	0.0012	0.0714	–1.86	8.97	–2.65	–3.87	–4.63	0	0	0
Adelaide Brighton Cement	abc1	–0.0006	0.0931	0.14	1.67	–2.96	–2.85	–2.94	0	0	0
Investa Property	wpr1	–0.0008	0.0375	–0.36	0.17	–3.27	–3.95	–6.30	0	0	0
Argo Investments	arg1	0.0028	0.0356	0.46	1.30	–4.44	–6.19	–5.02	0	0	0
Brickworks Ltd	bkw1	0.0056	0.0617	–0.01	1.73	–5.94	–6.74	–6.78	0	0	0
						100%	100%	100%	100%	100%	100%

The intuition of the PPI allows a plausible explanation. When the portfolio is benchmarked to the riskfree rate, Milton Corp is an attractive proposition. It has an average return well in excess of the riskfree rate (0.0120 per month), with relatively low standard deviation (0.0324). Since a portfolio heavily invested in this stock is likely to have a low probability of underperforming the benchmark, the PPI takes a substantial position in Milton Corp (43.97%). Doing so does not necessarily produce a portfolio with the highest Sharpe ratio (reward to total variability) or highest Sortino ratio (reward to downside variability), but it satisfies the objective function of maximising the decay rate in the underperformance probability.

In contrast, Milton Corp is less attractive as part of a portfolio benchmarked to the market index. Its average return in excess of the market is a modest 0.0068 per month, with a relatively high standard deviation. Intuitively, there is a non-trivial probability that a portfolio heavily invested in this stock will underperform the market. As a consequence, Stutzer's approach takes a smaller position in MLT1 (18.68%) and the optimal PPI-maximising weights are spread more evenly across the twenty stocks.

We conclude the discussion of optimal portfolio weights with a caution about estimation risk. The Sharpe ratio metric requires point estimates of the means and variances of candidate stocks, and their variance-covariance matrix. Recent evidence suggests that considerable estimation risk surrounds these sample estimators, so much so that the overall usefulness of the Sharpe ratio is questionable (Christie 2005, 2007).

Examining stock weightings tables 1 and 3, there are only modest differences between the weights assigned under the Sharpe-ratio, PPI, and Sortino-ratio approaches. The optimal weights under the Sharpe ratio are re-estimated using the robust covariance estimator of Ledoit and Wolf (2004). This shrinkage estimator corrects for biases that are likely to exist in a sample covariance matrix estimated in the standard fashion. The results (not explicitly reported) are qualitatively similar to those reported in tables 1 and 3. Nonetheless, in light of estimation risk, we caution against drawing strong inferences about differences in stock weightings under alternate approaches.

5.2 Fund Performance Measurement

The asymptotic (or 'large deviations') results behind the derivation of Stutzer's portfolio performance index require that the expected return on a fund in excess of the benchmark is positive. In other words, there is little point in calculating the decay rate in the underperformance probability if the expected excess return is negative. Accordingly, we restrict the ranking of funds to the subset of funds that, given sufficient time, would almost surely beat the chosen benchmark.

Following Foster and Stutzer (2004), each fund is subjected to a statistical test that the expected excess portfolio return equals zero (i.e. $E(\overline{R}_p) = 0$). This is a one-sided test against the alternative hypothesis that $E(\overline{R}_p) > 0$. Using a significance level of 5%, any fund for which the null hypothesis cannot be rejected is excluded from the subsequent fund ranking. Intuitively, investors are unlikely to be interested in funds for which there is no statistical evidence of benchmark outperformance.

Of the 295 funds examined, only 29 exhibit statistical support for outperformance.[10] Table 4 documents the distribution across fund categories of the population of Australian equity unit trusts and the subset of funds for which the PPI is calculated. Large blend and large growth funds dominate the population of Australian equity funds, with relatively few mid/small funds. However, it is predominantly the mid/small blend funds that exhibit statistical evidence of benchmark outperformance. Indeed, mid/small blend funds account for 59% of the subsample for which the PPI is calculated.

Table 4

Summary Statistics for Unit Trust Sample

Fund data are sourced from Morningstar Direct. The full sample comprises unit trusts with products categorised as Australian equity investments with monthly returns for the 60-month period between 2002–2006. From the full sample, the Portfolio Performance Index is calculated for the subset of funds with a statistically significant monthly return (in excess of the ASX200 market index).

Fund Categories	Full Sample	PPI ranked funds
General	1%	0%
Large blend	46%	24%
Large geared	1%	0%
Large growth	23%	0%
Large value	13%	7%
Mid/small blend	8%	59%
Mid/small growth	6%	7%
Mid/small value	2%	3%
Number of funds	295	29

Table 5 reports the identity of the 29 funds ranked in this study, along with the sample mean and standard deviation of their (log) excess return relative to the ASX200 index benchmark. This table allows comparison of the relative performance of the selected funds according to a number of metrics. First, an investor/analyst who is primarily concerned with end-of-period wealth will focus on the mean excess log return reported in column 4. The fund with the highest $\overline{R}_p$, known as the growth-optimal or log-optimal portfolio (Stutzer 2003; Hakannsson & Ziemba 1995), will generate the highest wealth asymptotically.[11] Foster and Stutzer (2004) highlight two important deficiencies in this metric. The growth-optimal portfolio is often heavily invested in high-risk assets, inducing

10. This ratio (approximately 10%) is very similar to that reported in Foster and Stutzer (2004, p. 17) for their sample of U.S. mutual funds. Such a small number of funds is consistent with evidence in the mutual fund literature that suggests that funds struggle to consistently outperform their chosen benchmark.

11. That is, the growth-optimal portfolio has the highest geometric growth rate in wealth. Equivalently, this portfolio has the highest log utility of wealth.

Table 5

Ranking of Fund Performance

The data are monthly fund returns from 2002–2006 drawn from Morningstar. The full sample comprises 295 unit trusts whose products are categorised as Australian equity investments. Only those funds with a statistically significant sample return in excess of the ASX200 index benchmark are included.

					Portfolio Performance Index		Sharpe Ratio	
Fund	Morningstar Category	Morningstar Ticker	$\bar{R}_p$	$\text{std}(\bar{R}_p)$	PPI	Rank	Sharpe	Rank
Suncorp Metway IMT Aust Equities Trust	large blend	4588	0.0020	0.0049	0.1044	1	0.4160	1
CFS FC Ws—Ws Small Companies Fund Core	mid/small blend	7006	0.0091	0.0220	0.0805	2	0.4122	2
AMP Capital Small Companies Fund	mid/small blend	10746	0.0074	0.0200	0.0687	3	0.3700	3
AMP Cap Small Companies Fund—C1 A	mid/small blend	8012	0.0073	0.0205	0.0632	4	0.3535	4
Prime Value Imputation Fund—Class B	large blend	13205	0.0106	0.0323	0.0613	5	0.3287	10
Prime Value Imputation Fund	large value	11233	0.0105	0.0324	0.0598	6	0.3248	11
Prime Value Growth Fund—Class B	large blend	13206	0.0075	0.0231	0.0568	7	0.3225	12
Prime Value Growth Fund	large blend	11232	0.0074	0.0233	0.0555	8	0.3193	13
GMO Australian Small Companies Trust	mid/small blend	8871	0.0062	0.0188	0.0550	9	0.3323	8
Challenger Premier Smaller Companies	mid/small blend	9115	0.0088	0.0267	0.0550	10	0.3291	9
BT PPSI—Ws Smaller Companies Fund	mid/small blend	6425	0.0074	0.0220	0.0541	11	0.3344	5
BT Wholesale—Smaller Companies Fund	mid/small blend	2725	0.0074	0.0220	0.0541	12	0.3344	6
BT Wholesale—Imputation Fund	large blend	5730	0.0027	0.0080	0.0515	13	0.3323	7
AMP FLI—AMP Small Companies	mid/small blend	1535	0.0064	0.0208	0.0486	14	0.3095	15
BT Investment—BT Smaller Companies Fund	mid/small blend	773	0.0068	0.0218	0.0469	15	0.3119	14
Macquarie—Small Companies Fund	mid/small blend	8602	0.0070	0.0262	0.0344	16	0.2669	16
Macquarie Master—Small Companies Fund	mid/small blend	5529	0.0069	0.0262	0.0339	17	0.2647	17
Credit Suisse—Australian Small Companies	mid/small growth	9257	0.0074	0.0295	0.0326	18	0.2512	18
Souls—Australian Equity Fund	mid/small blend	7893	0.0040	0.0162	0.0326	19	0.2443	19
Challenger Wholesale Australian Share Fund	large blend	3566	0.0016	0.0064	0.0306	20	0.2426	20
INVESCO Ws—Australian Smaller Companies	mid/small blend	2557	0.0056	0.0240	0.0274	21	0.2317	22
Aust Unity Acorn Capital Ws Microcap Trust	mid/small blend	7875	0.0084	0.0358	0.0273	22	0.2330	21
Investors Mutual Ws—Australian Smaller Companies	mid/small value	5340	0.0047	0.0206	0.0270	23	0.2310	23
Credit Suisse Priv—Australian Small Companies	mid/small growth	9258	0.0066	0.0295	0.0261	24	0.2251	26
Tyndall—Australian Share Wholesale Portfolio	large value	3987	0.0023	0.0105	0.0260	25	0.2219	27
Synergy (Inv) Perpetual Smaller Companies	mid/small blend	7565	0.0059	0.0255	0.0259	26	0.2307	24
Perpetual Wholesale Smaller Companies Fund	mid/small blend	4363	0.0055	0.0246	0.0246	27	0.2256	25
EQT Wholesale Small Companies Fund	mid/small blend	9121	0.0080	0.0365	0.0240	28	0.2197	28
BT Investment—BT Imputation Fund	large blend	5101	0.0018	0.0084	0.0229	29	0.2188	29

considerable downside risk. As such, choice of the growth-optimal portfolio ignores investor concerns over underperformance. Further, the growth-optimal portfolio is optimal asymptotically, yet Foster and Stutzer note that this long-run view may exceed the typical investor's horizon.

Nonetheless, it is interesting to compare fund rankings under the growth-optimal metric to the alternatives. For example, fund #9121 is ranked second last under the PPI and Sharpe ratio, yet it would be ranked highly under the growth-optimal metric—its average log excess return (0.0080) is four times greater than the highly-ranked fund #4588. Fund #4588 has a modest $\overline{R}_p$, yet it exhibits the fastest decay rate in underperformance probability and the highest Sharpe ratio.

A second metric reported in table 5 is the Sharpe ratio. An investor concerned with the average excess log return per unit of risk focuses on the Sharpe ratio in column 8. There is a reasonably high degree of consistency between the ranking of funds on the SR and the PPI. The top four funds are identical under these methods, as are the bottom two funds. However, non-trivial ranking differences can be observed. Fund #6425 is ranked fifth by the Sharpe ratio, but a modest eleventh under the PPI. Most ranking differences exist within the ranking range of 5 to 13. Indeed, the identity of the top-10 funds is notably different selection under the Sharpe ratio and PPI metrics.

Emphasising investor aversion to benchmark underperformance, the PPI calculates the decay rate in the underperformance probability. The funds in table 5 are ranked in descending order of their PPI. Decay rates range from 10.44% per month to 2.29% per month. Using the rule of 72, these estimates equate to the probability of underperformance halving every 7 months and 31 months respectively. For an investor aiming to minimise the probability of underperformance, these are useful (and simple) performance measures.

Note that the estimator of the performance index (5) assumes that fund returns are independent and identically distributed. To check the validity of this assumption, autocorrelations out to six lags are calculated for each fund. Only four of the 29 funds exhibit statistically significant autocorrelations. Nonetheless, Foster and Stutzer (2004) provide an alternate estimator for the case of non-IID fund returns. Re-estimating our PPI rankings using the robust estimator produces near-identical ranking of funds.[12]

6. Conclusions

The Australian funds-management industry is currently experiencing rapid expansion. As a result, there is an increasing level of interest in performance measurement, especially amongst unsophisticated investors. While the academic literature is rich with sophisticated risk-adjusted performance metrics, the appeal of these techniques is limited amongst industry professionals and their clients.

While Sharpe ratios and information ratios are widely employed in practice, these approaches to performance measurement are predicated on the assumption that stock and fund returns are normally distributed. The metrics also fail to

12. A Wilcoxon sign-rank test found no significant difference in the ranking of funds under the standard PPI and the robust estimator. This was the case irrespective of the choice of smoothing/bandwidth parameter K (K = 1, 3, 6). See Foster and Stutzer (2004, equation 25) for details of the robust estimator.

accommodate investor preferences for positive skewness (equivalently, downside risk aversion).

This paper explores the potential usefulness of the recently-proposed Portfolio Performance Index of Stutzer (2000). The PPI metric has a number of attractive features. By focusing on (the decay rate in) the probability that a portfolio will underperform a designated benchmark, the PPI captures the contemporary notion that investors view risk as the likelihood of not reaching investment targets. The intuition behind the PPI is straight forward and easily interpreted by professional and lay investors alike.

The empirical analysis presented in this paper centers around optimal portfolio construction and performance evaluation of Australian equity funds. Optimal stock weightings are shown to differ marginally under the PPI and Sharpe ratio metrics, thus providing an indication of the economic significance of return deviations from normality. Further, the characteristics of the resulting portfolios differ notably (especially the skewness of portfolio returns). Similarly, the ranking of equity funds displays non-trivial differences under the alternate metrics.

Given the substantial literature documenting departures from normality in stock returns, this paper's findings of differences in the composition of optimal portfolios and ranking of funds under alternate metrics is not surprising. Nonetheless, the results serve to both highlight the economic significance of return non-normality, and motivate further investigation of non-parametric techniques like Stutzer's PPI that closely map into investor objectives.

(Date of receipt of final transcript: August 6, 2007.
Accepted by Phil Dolan & Tom Smith, Special Issue Editors.)

References

Alles, L. & Spowart, A. 1995, 'Higher moments of Australian equities: Characteristics and determinants', *Accounting Research Journal,* vol. 8, pp 66–76.

Ang, A., Chen, J. & Xing, Y. 2006, 'Downside risk', *Review of Financial Studies,* vol. 19, pp. 1191–239.

Beedles, W.L. 1986, 'Asymmetry in Australian equity returns', *Australian Journal of Management,* vol. 11, pp. 1–12.

Beedles, W.L., Dodd, P. & Officer, RR. 1988, 'Regularities in Australian share returns', *Australian Journal of Management,* vol. 13, pp. 1–29.

Bodie, Z., Kane, A. & Marcus, A.J. 2005, *Investments*, sixth ed., McGraw-Hill Irwin, Boston.

Carhart, M.M. 1997, 'On persistence in mutual fund performance', *Journal of Finance,* vol. 52, pp. 57–82.

Christie, S. 2005, 'Is the Sharpe ratio useful in asset allocation?' working paper, Macquarie University, Sydney.

Christie, S. 2007, 'Beware the Sharpe ratio', working paper, Macquarie University, Sydney.

Daniel, K.M., Grinblatt, Titman, S. & Wermers, R. 1997, 'Measuring mutual fund performance with characteristic based benchmarks', *Journal of Finance,* vol. 52, pp. 1035–58.

Elton, E.J. & Gruber, M.J. 1995, *Modern Portfolio Theory and Investment Analysis*, fifth ed., Wiley, New York.

Ferson, W.E. & Schadt, R.W. 1996, 'Measuring fund strategy and performance in changing economic conditions', *Journal of Finance,* vol. 51, pp. 425–61.

Foster, F.D. & Stutzer, M. 2004, 'Performance and risk aversion of funds with benchmarks: A large deviations approach', working paper, Australian National University.

Goodwin, T.H. 1998, 'The information ratio', *Financial Analysts Journal,* vol. 54, pp. 34–43.

Gray, P., Kalotay, E. & McIvor, J. 1998, 'Testing the multivariate normality of Australian stock returns', *Australian Journal of Management,* vol. 23, pp. 135–50.

Hakannsson, N.H. & Ziemba, W.T. 1995, 'Capital growth theory' in *Handbook in Operational Research and Management Science: Finance,* vol. 9, eds. R.A. Jarrow, R.A. Maksimovic, & W.T. Ziemba, North-Holland, Amsterdam.

Henriksson, R.D. & Merton, R.C. 1981, 'On market timing and investment performance II: Statistical procedures for evaluating forecasting skills', *Journal of Business,* vol. 4, pp. 513–33.

Jensen, M.C. 1968,' The performance of mutual funds in the period 1945–1964', *Journal of Finance,* vol. 23, pp. 389–419.

Ledoit, O. & Wolf, M. 2004, 'Honey, I shrunk the sample covariance matrix', *Journal of Portfolio Management,* vol. 31, pp. 110–9.

Leland, H.E. 1999, 'Beyond mean-variances: Performance measurement in a nonsymmetrical world', *Financial Analysts Journal,* vol. 55, pp. 27–36.

Markowitz, H. 1952, 'Portfolio selection', *Journal of Finance,* vol. 7, pp. 77–91.

Markowitz, H. 1959, *Portfolio Selection: Efficient Diversification of Investments*, John Wiley, New York.

Praetz, P. & Wilson, E.J.G. 1978, 'The distribution of stock market returns: 1958–1973', *Australian Journal of Management,* vol. 3, pp. 79–9.

Sawicki, J. & Ong, F. 2000, 'Evaluating managed fund performance using conditional measures: Australian evidence', *Pacific-Basin Finance Journal,* vol. 8, pp. 505–28.

Sharpe, W.F. 1966, 'Mutual fund performance', *Journal of Business,* vol. 39, pp. 119–38.

Sharpe, W.F. 1994, 'The Sharpe ratio', *Journal of Portfolio Management,* vol. 21, pp. 49–58.

Sortino, F.A. & van der Meer, R. 1991, 'Downside risk', *Journal of Portfolio Management,* vol. 17, pp. 27–31.

Stutzer, M. 2000, 'A portfolio performance index', *Financial Analysts Journal,* vol. 56, pp. 52–61.

Stutzer, M.J. 2003, 'Portfolio choice with endogenous utility: A large deviations approach', *Journal of Econometrics,* vol. 116, pp. 365–86.

Treynor, J.L. & Mazuy, K.K. 1966, 'Can mutual funds outguess the market?', *Harvard Business Review,* July-August, pp. 131–6.