

Article

Sensitivity of Performance Indexes to Disaster Risk

Jiro Hodoshima *  and Toshiyuki Yamawake * 

Faculty of Economics, Nagoya University of Commerce and Business, 4-4 Sagamine, Komenoki-cho, Nisshin-shi, Aichi 470-0193, Japan

* Correspondence: hodoshima@nucba.ac.jp (J.H.); yamawake@nucba.ac.jp (T.Y.);
Tel.: +81-561-73-2111 (J.H. & T.Y.); Fax: +81-561-73-1202 (J.H. & T.Y.)

Abstract: We examine how sensitive the new performance indexes incorporating high moments and disaster risk are to disaster risk. The new performance indexes incorporating high moments and disaster risk are the Aumann-Serrano performance index and Foster-Hart performance index proposed by Kadan and Liu. These performance indexes provide evaluations sensitive to the underlying risk. We show, by numerical examples and empirical examples, how sensitive these indexes are to disaster risk. Although these indexes are known to be either quite sensitive or excessively sensitive to disaster risk or maximum loss in the literature, we show by the regression analysis of the index and summary statistics these indexes are in fact not excessively sensitive to maximum loss in representative stock data, which contain disastrous observations. The numerical estimate of the Foster-Hart performance index is found to be effective in showing the performance index. Our analysis suggests these indexes can handle various empirical data containing quite disastrous observations.



Citation: Hodoshima, Jiro, and Toshiyuki Yamawake. 2021. Sensitivity of Performance Indexes to Disaster Risk. *Risks* 9: 40. <https://doi.org/10.3390/risks9020040>

Academic Editor: Mogens Steffensen

Received: 26 January 2021

Accepted: 8 February 2021

Published: 13 February 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: Aumann-Serrano performance index; Foster-Hart performance index; Sharpe ratio; maximum loss

JEL Classification: G11; C22; C46

1. Introduction

Recently, new performance measures incorporating high moments and disaster risk are proposed by Kadan and Liu (2014). The new performance measures are reciprocals of the risk indexes proposed by Aumann and Serrano (hereafter AS) (2008) and Foster and Hart (hereafter FH) (2009) based on axiomatic approaches. Kadan and Liu (2014) demonstrated the use of the new performance indexes is useful to provide risk-averse assessments of various assets compared to the de facto industry standard performance measure, i.e., the Sharpe ratio. These conservative assessments shed new light to provide very different evaluations compared to the Sharpe ratio, which incorporates only the first two moments. There are many performance measures proposed in the literature. However, most of those performance measures proposed before were put forward in a rather ad hoc way and not based on economic principles. See, e.g., articles such as Cogneau and Hubner (2009a, 2009b), Eling and Schuhmacher (2007), and Farinelli et al. (2008) for various performance measures proposed in the literature. On the other hand, the AS and FH performance measures are based on axiomatic approaches and different from those ad hoc performance measures. However, there are no further studies of the properties of the new performance measures since Kadan and Liu (2014) except Hodoshima (2019), who only studies the property of the AS performance measure¹. The AS performance measure is known to

¹ Hodoshima (2019) studied an index called the inner rate of risk aversion (IRRA) proposed by Miyahara (2010, 2014) obtained from utility indifference pricing when the underlying utility function is an exponential one. However, the IRRA turns out to be equivalent to the AS performance index (cf., Hodoshima and Miyahara 2020).

be quite sensitive to the underlying risk of targets in question (cf., [Kadan and Liu 2014](#); [Miyahara 2014](#); [Ban et al. 2016](#); [Hodoshima 2019](#)). The FH performance measure is also known to be extremely sensitive to rare disasters or maximal loss (cf., [Foster and Hart 2009](#); [Kadan and Liu 2014](#); [Anand et al. 2016, 2017](#); [Riedel and Hellmann 2015](#)). We address the issue of sensitivity of the AS and FH performance measures to disaster risk in this study. It is not clear in the existing literature how sensitive these new performance measures are to disaster risk. The AS and FH performance measures correspond to the AS and FH risk measures, one-to-one. Although we focus on the AS and FH performance measures in this study, therefore, our findings in this study can be equally applied to the AS and FH risk measures in the opposite sense that good (bad) performance is equivalent to being less (more) risky.

We use numerical examples and empirical examples to see how sensitive the new performance measures are to disaster risk or maximal loss. We present numerical examples of several performance measures to show characteristics of the new performance measures. The performance measures we study in the numerical examples are the AS and FH performance measures, Sharpe ratio, Sortino ratio, and Calmar ratio. By comparing these performance measures numerically, we can see more clearly the merits and demerits of the AS and FH performance measures as compared to the more traditional performance measures proposed before. We also study empirical examples to see how the new performance measures are sensitive to disaster risk or maximal loss. As empirical examples we use the DOW 30 stocks, which make up the world-leading Dow Jones Industrial Average Index (DOW). We consider the DOW 30 stocks are representative stocks. When we show daily and monthly data of the DOW 30 stocks as empirical examples in this study, we can see they contain quite disastrous days or months, which looks more disastrous than our numerical examples where the FH performance measure attains a lower limit of zero. However, we show both the AS and FH performance measures can handle these data to produce sensible scores in our empirical examples. We show the new performance measures, particularly the FH performance measure, are, unlike the previous predominant view in the literature, able to produce robust scores to real data containing quite disastrous observations. The numerical estimate of the FH performance measure by grid search is found to be effective in showing a good estimate of the FH performance measure. We also show by the regression analysis of the new performance measure and summary statistics including percentiles how the new performance measures are sensitive to disaster risk or maximal loss in typical stock data.

In the following, we review the existing literature with respect to sensitivity of the AS and FH performance measures to disaster risk. In their seminal paper, [Aumann and Serrano \(2008\)](#) showed the AS risk index satisfies first- and second-order stochastic dominance when the underlying gamble is restricted to take finite values and noted the AS risk index is sensitive to losses. [Miyahara \(2010, 2014\)](#) showed the IRRA, equivalent to the AS performance index, is sensitive to losses by numerical examples. [Kadan and Liu \(2014\)](#) described the reciprocal of the AS risk index can be used as a performance index and demonstrated through several empirical examples it can be quite useful to provide risk-sensitive evaluation compared to the Sharpe ratio. [Ban et al. \(2016\)](#) showed the sensitivity of the IRRA to disaster risk by numerical examples. [Hodoshima \(2019\)](#) examined the properties of the IRRA using examples of normal mixture distributions.

On the other hand, [Foster and Hart \(2009\)](#) proved the FH risk index also satisfies first- and second-order stochastic dominance when the underlying gamble is restricted to take finite values and remarked it is quite sensitive to the maximal loss as an implication of their proposition, i.e., Proposition 3 of [Foster and Hart \(2009\)](#). [Foster and Hart \(2009\)](#) showed there exists a unique positive solution $R(g)$ of the implicit equation

$$E \left[\ln \left(1 + \frac{g}{R(g)} \right) \right] = 0$$

for a gamble g when a gamble g takes finite values with $E[g] > 0$ and $P(g < 0) > 0$ where $R(g)$ denotes the FH risk measure. Kadan and Liu (2014) also remarked the FH performance measure is extremely sensitive to the maximal loss by considering a composite gamble with a small probability of a large loss, i.e., Proposition 8 of Kadan and Liu (2014). Anand et al. (2016) and Anand et al. (2017) use as the FH risk measure the extended FH risk given by Riedel and Hellmann (2015) for gambles with continuous distribution. In other words, Riedel and Hellmann (2015) extended the FH risk measure for a gamble with continuous distribution by the maximum loss L , i.e.,

$$R(g) = \begin{cases} r^* & \text{such that } E[\ln(1 + \frac{g}{r^*})] = 0 \text{ if } E[\ln(1 + \frac{g}{L})] < 0 \\ L & \text{if } E[\ln(1 + \frac{g}{L})] \geq 0. \end{cases}$$

Riedel and Hellmann (2015) remarked the FH risk measure does not exist for many common distributions and that the sign of the expectation $E[\ln(1 + \frac{g}{L})]$ determines whether the FH risk measure exists or not. In other words, the FH risk measure defined by Foster and Hart (2009) does (not) exist when the sign of $E[\ln(1 + \frac{g}{L})]$ is negative (nonnegative), and the maximal loss L is used as the extended FH risk measure when the FH risk measure does not exist. Since Riedel and Hellmann (2015), the FH risk measure have been identified in some cases as the maximum loss. However, to the best of our knowledge no study exists to examine if the maximum loss or the reciprocal of the maximum loss of a gamble is appropriate in real data. In this study, we verify this thesis empirically using the numerical estimate of the FH performance measure proposed by Kadan and Liu (2014) and the reciprocal of the maximum loss with the help of the sign of the expectation $E[\ln(1 + \frac{g}{L})]$. As a result, we show the numerical estimate of the FH performance measure by grid search is effective in finding the FH performance measure or the extended FH performance measure defined by the reciprocal of the maximal loss.

The rest of the article is organized as follows. In Section 2, we describe methods, i.e., the AS and FH performance measures as well as the extended FH performance measure. In Section 3, we provide results, i.e., numerical examples of a set of performance measures, including the AS and FH performance measures, and empirical examples. In Section 4, we provide discussion of results. In Section 5, we present concluding comments.

2. Methods (the AS and FH Performance MEASURE and FH Extended Performance Measure)

In this section, we describe the AS and FH performance measures and FH extended performance measure. We follow Kadan and Liu (2014) to describe the AS and FH performance measures, i.e., Proposition 1 and 2 of Kadan and Liu (2014).

First we describe the AS performance measure. We begin to give the following two definitions.

Definition 1. A gamble g is wealth-uniformly rejected by an investor with utility function u , if u rejects g at all initial wealth levels.

Definition 2. A gamble g wealth-uniformly dominates gamble g' if whenever g is wealth-uniformly rejected by a utility function u , g' is also wealth-uniformly rejected by u .

Based on the above definitions, the following result can be obtained.

Proposition 1 (Aumann and Serrano 2008; Hart 2011). Wealth-uniform dominance induces a complete order on the set \mathcal{G} of gambles that extends second-order stochastic dominance. This order can be represented by a performance index $P^{AS}(g)$ assigned to any gamble $g \in \mathcal{G}$, which is given by the unique positive solution to the implicit equation

$$E[\exp(-P^{AS}(g)g)] = 1.$$

That is, for any two gambles g and g' , g wealth-uniformly dominates g' if and only if $P^{AS}(g) \geq P^{AS}(g')$. The set \mathcal{G} denotes the set of gambles.

Next, we describe the FH performance measure. We assume the class of utility functions \mathcal{U}^* satisfy the three conditions; (i) decreasing absolute risk aversion, (ii) increasing relative risk aversion, (iii) $\lim_{w \downarrow 0} u(w) = -\infty$.

Definition 3. A gamble g is utility-uniformly rejected at an initial wealth level w_0 if all utility functions $u \in \mathcal{U}^*$ reject g at w_0 .

Definition 4. A gamble g utility-uniformly dominates gamble g' if whenever g is utility-uniformly rejected at an initial wealth level w_0 , g' is also utility-uniformly rejected at w_0 .

Then we have the following result.

Proposition 2 (Foster and Hart 2009; Hart 2011). Utility-uniform dominance induces a complete order on \mathcal{G} that extends second-order stochastic dominance. This order can be represented by a performance index $P^{FH}(g)$ assigned to any gamble $g \in \mathcal{G}$, which is given by the unique positive solution to the implicit equation

$$E[\ln(1 + P^{FH}(g)g)] = 0.$$

That is, for any two gambles g and g' , g utility-uniformly dominates g' if and only if $P^{FH}(g) \geq P^{FH}(g')$.

Therefore, the AS performance index of a gamble g is defined to be a positive solution α of the implicit equation given by

$$E[\exp(-\alpha g)] = 1.$$

On the other hand, the FH performance index of a gamble g is defined to be a positive solution γ of the implicit equation

$$E[\ln(1 + \gamma g)] = 0.$$

The AS and FH risk indexes are the reciprocals of the AS and FH performance indexes, i.e., the AS risk index proposed by Aumann and Serrano (2008) is given by $1/\alpha$ and the FH risk index proposed by Foster and Hart (2009) is given by $1/\gamma$. Hodoshima and Miyahara (2020) extended the AS performance index to include the negative solution of the implicit equation of the AS performance index and provided a sufficient condition for the existence of the unique negative solution of the implicit equation.

Riedel and Hellmann (2015) gave an interesting theorem regarding a gamble with continuous distribution by introducing finite gambles that approximate the continuous gamble.

Proposition 3 (Riedel and Hellmann 2015). Let g be a gamble with maximal loss $L > 0$. Let g_n be a sequence of finite-valued gambles with $g_n \uparrow g$ a.s., where each g_n has the same maximal loss L . Denote by $\rho_n \equiv \rho(g_n) > L$ their FH risk measure. Then the following statements hold true:

1. The sequence ρ_n is decreasing. Put $\rho_\infty = \lim \rho_n \geq L$ for its limit.
2. If $E[\ln(1 + g/L)] < 0$, then ρ_∞ is the unique positive solution of the FH implicit equation $E[\ln(1 + g/\gamma)] = 0$.
3. If $E[\ln(1 + g/L)] \geq 0$, then the Foster-Hart equation has no solution and $\rho_\infty = L(g)$.

In other words, Riedel and Hellmann (2015) defined the maximal loss L as the extended FH risk measure when $E[\ln(1 + g/L)]$ is nonnegative, i.e., when the FH risk measure does not exist.

3. Results

In this section, we present numerical examples of a set of performance measures including the AS and FH performance measures and empirical examples of the DOW 30 stocks.

3.1. Numerical Examples of the Performance Measures

We first present numerical examples of a set of performance measures using random variables of cash flows. Although Kadan and Liu (2014) presented empirical examples of the AS and FH performance measures in various financial assets, the characteristics of the two performance measures were not fully explained there. We intend to show the characteristics of the two performance measures in comparison to the traditional performance measures of the Sharpe ratio, Sortino ratio, and Calmar ratio. The Sortino ratio and Calmar ratio are derived from the Sharpe ratio by replacing the standard deviation in the Sharpe ratio by other risk measures. The Sortino ratio replaces standard deviation in the Sharpe ratio by the downside deviation in order to take into account only downside risk instead of both downside and upside risk in standard deviation. The Calmar ratio replaces standard deviation in the Sharpe ratio by another risk measure of the maximum drawdown where a drawdown is a peak-to-trough decline. Comparing the two performance measures with these traditional performance measures, we can see how the new performance measures function relative to the old performance measures.

We consider the following two sets of random variables in Example 1 and 2 given in Tables 1 and 2. We consider four cases of random variables in Example 1 where values each random variable takes are given with corresponding probabilities where pr stands for probability. Mean, s.d., third, and fourth, skew, kurt, downrisk, and maxdrawd are respectively mean, standard deviation, the third central moment, the fourth central moment, skewness, kurtosis, downside deviation, and maximum drawdown for each random variable. Similar numerical examples were examined in Hodoshima (2020a) where the traditional performance measures were compared with the AS performance measure.

Sharpe, Sortino, Calmar, AS, and FH denote respectively the Sharpe ratio, Sortino ratio, Calmar ratio, AS performance measure, and FH performance measure.

We estimate the two performance measures by the generalized method of moments (GMM) estimator, as described in Kadan and Liu (2014). In particular, we find the two performance measures via grid search for the solutions of the sample analogs of the implicit equations for the two performance measures. The GMM estimator is consistent and asymptotically normally distributed. The implicit equations of the two performance measures are given as follows. The AS performance measure of a gamble g is given by α , which is a unique solution α of the implicit equation

$$E[\exp(-\alpha g)] = 1. \quad (1)$$

On the other hand, the FH performance index of a gamble g is defined by γ , which is the unique solution of the implicit equation

$$E[\ln(1 + \gamma g)] = 0. \quad (2)$$

In Example 1, the case with a higher number dominates the case with a lower number since the former is larger than the latter with probability one. Hence, appropriate performance measures ought to take higher values in the case with a higher number than the case with a lower number, which is the property called *monotonicity*. The de facto industry standard performance measure, the Sharpe ratio, fails to provide a larger value in the case with a higher number than the case with a lower number. Therefore, the Sharpe ratio does not satisfy monotonicity, one of the most fundamental criteria for performance measures. On the other hand, the Sortino ratio and Calmar ratio both satisfy this criterion. However, the Calmar ratio increases very much in the case with a higher number where a large value replaces a small value in the case with a lower number with a small probability 0.009.

Therefore, the Calmar ratio is too sensitive to the increase of a value in the random variable with a small probability 0.009 in the case with a higher number. The Sortino ratio increases gradually in the case with a higher number so that it satisfies monotonicity in Example 1. The AS performance measure does satisfy to provide a larger number in the case with a higher number than the case with a lower number². We remark the AS performance measures in case 2–4 are the same up to the third decimal point in Example 1 but that the AS performance measure in the case with a higher number is in fact larger in four decimal places or less than the case with a lower number in case 2–4 in Example 1. The increase of the AS performance measure in the case with a higher number is very small and hence not sensitive to gains of the underlying random variable. The FH performance measure increases more clearly than the AS performance measure in the case with a higher number than the case with a lower number. This implies the FH performance measure is more sensitive to gains than the AS performance measure, which has never been mentioned in the literature. The magnitude of the FH performance measure is similar to that of the AS performance measure and Sortino ratio.

Table 1. Example 1 of Four Cases of Cash Flows.

	Case 1		Case 2		Case 3		Case 4	
	value	pr	value	pr	value	pr	value	pr
	−1	0.4	−1	0.4	−1	0.4	−1	0.4
	2	0.6	2	0.591	2	0.591	2	0.591
			5	0.009	10	0.009	20	0.009
mean	0.8		0.827		0.872		0.962	
s.d.	1.470		1.518		1.704		2.332	
third	−1.296		−0.831		5.069		59.742	
fourth	5.443		8.305		68.350		1188.914	
skew	−0.408		−0.238		1.025		4.710	
kurt	1.167		1.563		8.107		40.196	
downrisk	1.138		1.155		1.184		1.241	
maxdrawd	1.5		1.2		1.1		1.05	
Sharpe	0.544		0.545		0.512		0.413	
Sortino	0.703		0.716		0.737		0.775	
Calmar	0.533		0.689		0.793		0.916	
AS	0.782		0.785		0.785		0.785	
FH	0.746		0.751		0.755		0.760	

Example 1 provides four cases of random variables of uncertain cash flows where values each random variable takes are given with corresponding probabilities where pr stands for probability. Mean, s.d., third, and fourth, skew, kurt, downrisk, and maxdrawd are respectively mean, standard deviation, the third central moment, the fourth central moment, skewness, kurtosis, downside risk, and maximum drawdown for each random variable. Sharpe, Sortino, Calmar, AS, and FH denote respectively the Sharpe ratio, Sortino ratio, Calmar ratio, AS performance measure, and FH performance.

In Example 2, a random variable in each case has a disaster −15 with a small probability 0.001 and other cashflows with remaining probabilities proportional to probabilities in each case in Example 1. The disaster risk −15 with a small probability 0.001 does not affect much the Sharpe ratio and Sortino ratio but does affect significantly the Calmar ratio. However, the Calmar ratio is again seen to be too sensitive to the increase of a value in the random variable with a small probability 0.009×0.999 in the case with a higher number. Hence it is not a reliable performance measure. The AS performance measure becomes less than half in Example 2 than in Example 1. We can say the disaster risk has a large effect on

² In fact, the AS performance measure was proved to satisfy monotonicity (cf., [Aumann and Serrano 2008](#); [Hodoshima and Miyahara 2020](#)).

the AS performance measure. Thus, the AS performance measure is sensitive to losses or the maximal loss of the underlying random variable. The FH performance measure has an even more significant negative effect by the disaster risk. In Table 2, 0^+ denotes limit on the right, i.e., the limit of sequences of positive numbers converging to zero. It becomes virtually zero, which is a lower bound of the FH performance measure, in every case in Example 2. Adding larger positive values in the case with a higher number does not affect the FH performance measure in Example 2. Therefore, the FH performance measure is virtually determined by disaster risk or maximal loss. Hence, it is the most sensitive measure to disaster risk, which conforms to the previous studies of the FH performance measure (cf., Foster and Hart 2009; Kadan and Liu 2014; Anand et al. 2016, 2017; Riedel and Hellmann 2015). Hence, the newly introduced performance measures of the AS and FH performance measures are both quite sensitive to losses of the underlying random variable.

Table 2. Example 2 of Four Cases of Cash Flows.

	Case 1		Case 2		Case 3		Case 4	
	value	pr	value	pr	value	pr	value	pr
	−15	0.001	−15	0.001	−15	0.001	−15	0.001
	−1	a	−1	a	−1	a	−1	a
	2	b	2	c	2	c	2	c
			5	d	10	d	20	d
mean	0.784		0.811		0.856		0.946	
s.d.	1.552		1.598		1.775		2.385	
third	−5.125		−4.674		1.216		55.888	
fourth	67.430		70.744		131.818		1256.200	
skew	−1.372		−1.146		0.217		4.120	
kurt	11.636		10.854		13.265		38.832	
downrisk	1.233		1.249		1.276		1.330	
maxdrawd	8.5		4		2.5		1.75	
Sharpe	0.505		0.508		0.482		0.397	
Sortino	0.636		0.649		0.671		0.712	
Calmar	0.092		0.203		0.342		0.541	
AS	0.327		0.328		0.329		0.329	
FH	0^+		0^+		0^+		0^+	

Example 2 provides four cases of random variables of uncertain cash flows where values each random variable takes are given with corresponding probabilities where pr stands for probability. Mean, s.d., third, and fourth, skew, kurt, downrisk, and maxdrawd are respectively mean, standard deviation, the third central moment, the fourth central moment, skewness, kurtosis, downside risk, and maximum drawdown for each random variable. In the table, a, b, c, and d denote respectively the probability given by 0.4×0.999 , 0.6×0.999 , 0.591×0.999 , and 0.009×0.999 . Sharpe, Sortino, Calmar, AS, and FH denote respectively the Sharpe ratio, Sortino ratio, Calmar ratio, AS performance measure, and FH performance measure. In the table, 0^+ denotes limit on the right, i.e., the limit of sequences of positive numbers converging to zero.

We introduce another example, Example 3 given in Table 3, to show the problematic nature of the traditional performance measures of the Sharpe ratio, Sortino ratio, and Calmar ratio. Example 3 has a loss of −5, which is not huge but sizable compared to the disaster risk −15 in Example 2, with a small probability 0.001. It has four cases where a case with a higher number is larger than a case with a smaller number with probability 1. The loss −5 produces larger skewness and kurtosis in absolute value in Example 3 than in Example 2. We can see absurd values of the Sharpe ratio in Example 3, failing to satisfy monotonicity again. In Example 3, the Sortino ratio also fails to satisfy monotonicity. Therefore, we cannot always trust the Sortino ratio because it does give irrational values, depending on the underlying random variable. We can also observe the Calmar ratio is

again too sensitive to the increase of a value in the random variable with small probability 0.009. Hence, we cannot trust the Calmar ratio as an appropriate performance measure. In Example 3, the FH performance index is again 0^+ , the limit of sequences of positive numbers converging to zero, in every case of Example 3. Adding larger positive values does not change the FH performance measure in Example 3. Hence, the FH performance measure is again virtually determined by the disaster risk of -5 in Example 3. One may say this indicates that the FH performance measure is excessively sensitive to disaster risk. On the other hand, the AS performance measure provides larger scores in Example 3 than in Examples 1 and 2. This also conforms to the performance of mean in the three examples, i.e., mean in Example 3 is larger than in Example 1 and 2. Therefore, the AS performance measure is sensitive to losses of the underlying random variable but not excessively sensitive to disaster risk as in the FH performance measure. We can see the AS performance measure is again insensitive to gains in Example 3. One may consider the AS performance measure, sensitive to losses but not excessively sensitive to disaster risk, is more appropriate than the FH performance measure since it can provide assessments more often.

Table 3. Example 3 of Four Cases of Cash Flows.

	Case 1		Case 2		Case 3		Case 4	
	value	pr	value	pr	value	pr	value	pr
	−5	0.001	−5	0.001	−5	0.001	−5	0.001
	1	0.999	1	0.99	1	0.99	1	0.99
			5	0.009	10	0.009	20	0.009
mean	0.994		1.03		1.075		1.165	
s.d.	0.190		0.423		0.871		1.805	
third	−0.215		0.344		6.174		59.898	
fourth	1.291		3.558		58.467		1134.119	
skew	−31.575		4.537		9.330		10.187	
kurt	998.001		110.914		101.391		106.860	
downrisk	0.190		0.193		0.206		0.255	
maxdrawd	6		2		1.5		1.25	
Sharpe	5.241		2.434		1.234		0.645	
Sortino	5.244		5.337		5.216		4.571	
Calmar	0.166		0.515		0.717		0.932	
AS	1.319		1.320		1.320		1.320	
FH	0^+		0^+		0^+		0^+	

Example 3 provides four cases of random variables of uncertain cash flows where values each random variable takes are given with corresponding probabilities where pr stands for probability. Mean, s.d., third, and fourth, skew, kurt, downrisk, and maxdrawd are respectively mean, standard deviation, the third central moment, the fourth central moment, skewness, kurtosis, downside risk, and maximum drawdown for each random variable. Sharpe, Sortino, Calmar, AS, and FH denote respectively the Sharpe ratio, Sortino ratio, Calmar ratio, AS performance measure, and FH performance measure. In the table, 0^+ denotes limit on the right, i.e., the limit of sequences of positive numbers converging to zero.

Although our numerical examples are limited, we can summarize our numerical comparisons as follows. Overall, the traditional performance measures of the Sharpe ratio, Sortino ratio, and Calmar ratio are not reliable since they do not either satisfy monotonicity or sometimes give irrational evaluations depending on the underlying target in question. On the other hand, the AS and FH performance measures are reliable, i.e., satisfy monotonicity, when they are well defined. However, the FH performance measure is excessively sensitive to the maximal loss, which makes the FH performance measure incapable of providing appropriate assessments. On the other hand, the AS performance

measure can provide assessments more often than the FH performance measure. The AS performance measure is quite sensitive to losses but not excessively sensitive to the maximal loss as in the FH performance measure. The AS performance measure is less sensitive to gains than the FH performance measure when the two performance measures both can provide assessments.

In the next subsection, we provide empirical examples to show how the AS and FH performance measures function when we can compute the two performance measures. In particular, we show by empirical examples how the FH performance measure performs when we can obtain its assessment, which we could observe only once in Example 1 in three numerical examples in this section.

3.2. Empirical Results

In this subsection, we present empirical results of evaluations by the two new performance measures and the Sharpe ratio for the DOW 30 components³ as of 2 April 2019. The DOW 30 components are listed in Table 4. Our sample period for daily (monthly) return data is from 4 January 2005 (February 2005) till 30 December 2019 (December 2019)⁴. As stock returns, we use log-returns in this paper. The same data were studied by [Hodoshima and Yamawake \(2020\)](#) where only winners and losers of the DOW 30 components were described without percentiles including the maximal loss. On the other hand, the current study focuses on the issue of sensitivity of the new performance measures to disaster risk. We cursorily describe performance of the DOW 30 stocks in this study. The list of the DOW 30 components is given in Table 4.

Table 4. List of DOW 30 Components (as of 2 April 2019).

3M (MMM)	American Express (AXP)	Apple (AAPL)
Boeing (BA)	Caterpillar (CAT)	Chevron (CVX)
Cisco (CSCO)	Coca-Cola (KO)	Disney (DIS)
Dow Chemical (DOW)	Exxon Mobil (XOM)	Goldman Sachs (GS)
Home Depot (HD)	IBM (IBM)	Intel (INTC)
Johnson & Johnson (JNJ)	JP Morgan (JPM)	McDonald's (MCD)
Merck (MRK)	Microsoft (MSFT)	Nike (NKE)
Pfizer (PFE)	Procter & Gamble (PG)	Travelers (TRV)
United Technologies (UTX)	UnitedHealth (UNH)	Verizon (VZ)
Wal-Mart (WMT)	Walgreens (WBA)	Visa (V)

We present summary statistics of daily return data in Table 5 and percentiles, including the maximum and minimum, in Table 6. In the tables, s.d., max, min, 80%, 60%, 40%, and 20% denote respectively standard deviation, maximum, minimum, 80% percentile, 60% percentile, 40% percentile, and 20% percentile. Mean in the last row in Tables 5 and 6 denotes mean of summary statistics and percentiles over 29 stocks. Mean ranges from 0.017 in Walgreens to 0.114 in Apple. Standard deviation ranges from 1.017 in Johnson & Johnson to 2.350 in JP Morgan. Skewness shows 11 stocks are negatively skewed and that 18 stocks are positively skewed. Kurtosis shows all the data have heavy tails compared to the normal distribution. The maximum ranges from 8.975 in McDonald's to 29.829 in UnitedHealth. The minimum ranges from −8.226 in Procter & Gamble to −23.228 in JP Morgan. These minimum values are for daily returns and hence considered to be quite disastrous losses. Four other percentiles are given in Table 6. They are listed from larger values (80% percentile) to smaller values (20% percentile).

³ We drop Dow Chemical (DOW) in our analysis because Dow Chemical was born as a new company, separated from DowDuPont in April 2019. We include Visa even though its sample period begins from March 19, 2008 since it was listed from 19 March 2008.

⁴ The only exception is Visa, where the sample period for daily (monthly) return data is from 20 March 2008 (April 2008) till December 30 2019 (December 2019).

Table 7 presents the three performance measures, the AS and FH performance measures and the Sharpe ratio, for daily return data. We do not provide the Sortino ratio and Calmar ratio in this section since our focus is on the two performance measures and the Sharpe ratio is the de facto industry standard performance measure to compare. Mean in the last row in Table 7 denotes mean of performance measures over 29 stocks. We can obtain the AS performance measure in every stock of the DOW 30 stocks. In other words, we can obtain the AS performance measure without much difficulty in representative real stock data. Therefore, obtaining the AS performance measure is not a problem in our daily stock data. The AS performance measure ranges from 0.010 in Goldman Sachs to 0.086 in McDonald's. Therefore, the AS performance measure scores are much smaller than those in the numerical examples in the previous subsection although the maximum loss is in some stocks larger than the numerical examples in the previous section. The difference between the AS performance measure and Sharpe ratio is large in outperforming stocks but small in underperforming stocks.

Table 5. Summary Statistics of Daily Return Data of DOW 30 Stocks.

Name	Mean	s.d.	Skew	Kurt
MMM	0.030	1.362	−0.650	11.578
AXP	0.031	2.149	0.016	17.282
AAPL	0.114	2.037	−0.297	9.008
BA	0.059	1.742	0.009	7.651
CAT	0.041	1.972	−0.222	8.652
CVX	0.037	1.608	0.056	15.706
CSCO	0.031	1.777	−0.352	14.852
KO	0.044	1.082	0.262	16.402
DIS	0.049	1.587	0.231	12.239
XOM	0.020	1.475	0.009	16.475
GS	0.026	2.224	0.262	20.094
HD	0.053	1.574	0.324	8.621
IBM	0.018	1.348	−0.268	9.470
INTC	0.036	1.785	−0.225	8.496
JNJ	0.034	1.017	−0.117	15.889
JPM	0.044	2.350	0.331	21.752
MCD	0.060	1.184	0.154	8.245
MRK	0.043	1.547	−0.311	15.233
MSFT	0.056	1.630	0.062	13.307
NKE	0.078	1.667	0.314	11.062
PFE	0.026	1.365	−0.162	10.394
PG	0.033	1.070	−0.089	10.619
TRV	0.044	1.651	0.243	29.981
UTX	0.037	1.422	0.093	10.108
UNH	0.055	1.941	0.429	29.617
VZ	0.033	1.277	0.234	11.338
V	0.090	1.838	0.012	11.664
WMT	0.030	1.207	0.069	14.269
WBA	0.017	1.663	−0.436	14.073
mean	0.044	1.605	−0.001	13.934

In the table, s.d., skew, and kurt denote respectively standard deviation, skewness, and kurtosis.

Table 6. Percentiles of Daily Return Data of DOW 30 Stocks.

Name	Max	Min	80%	60%	40%	20%
MMM	9.420	−13.863	0.884	0.285	−0.147	−0.726
AXP	18.771	−19.352	1.083	0.341	−0.209	−0.943
AAPL	13.019	−19.747	1.438	0.498	−0.230	−1.122
BA	14.378	−9.353	1.219	0.423	−0.255	−1.047
CAT	13.735	−15.686	1.329	0.411	−0.269	−1.160
CVX	18.942	−13.341	1.094	0.380	−0.227	−0.956
CSCO	14.799	−17.687	1.122	0.344	−0.212	−0.971
KO	12.997	−9.068	0.710	0.253	−0.129	−0.604
DIS	14.818	−10.231	1.029	0.322	−0.202	−0.909
XOM	15.863	−15.027	0.969	0.297	−0.222	−0.865
GS	23.482	−21.022	1.307	0.386	−0.281	−1.176
HD	13.161	−8.579	1.038	0.325	−0.200	−0.921
IBM	10.899	−8.662	0.903	0.278	−0.203	−0.780
INTC	11.199	−13.221	1.215	0.406	−0.257	−1.098
JNJ	11.537	−10.578	0.689	0.215	−0.132	−0.571
JPM	22.392	−23.228	1.144	0.328	−0.249	−1.049
MCD	8.975	−8.316	0.822	0.301	−0.133	−0.713
MRK	12.251	−15.944	0.990	0.316	−0.213	−0.831
MSFT	17.063	−12.458	1.042	0.319	−0.205	−0.887
NKE	11.876	−12.596	1.081	0.372	−0.217	−0.946
PFE	9.687	−11.232	0.909	0.270	−0.231	−0.806
PG	9.726	−8.226	0.708	0.224	−0.154	−0.612
TRV	22.758	−20.067	0.898	0.303	−0.162	−0.806
UTX	12.793	−9.170	0.965	0.308	−0.184	−0.832
UNH	29.829	−20.624	1.151	0.392	−0.227	−1.031
VZ	13.656	−8.413	0.861	0.300	−0.176	−0.792
V	13.974	−14.669	1.175	0.411	−0.184	−1.014
WMT	10.502	−10.740	0.785	0.261	−0.164	−0.723
WBA	15.388	−16.237	1.095	0.306	−0.261	−1.002
mean	14.755	−13.701	1.0226	0.330	−0.205	−0.893

In the table, max, min, 80%, 60%, 40%, and 20% denote respectively maximum, minimum, 80% percentile, 60% percentile, 40% percentile, and 20% percentile.

The FH performance measure is generally similar to the AS performance measure. Example 1 in the previous subsection shows the FH performance measure is similar to the AS performance measure when there is not disaster risk, which can be the reason why the FH performance measure is similar to the AS performance measure, although the minimum values in Table 6 show the existence of severe negative returns in many stocks. The existence of these severe negative returns seems to have downward effects on the FH performance measure as well as the AS performance measure in daily data. The two performance measures are substantially small compared to those in Example 1 and the AS performance measure in Example 2 and 3 at the previous subsection.

Table 7. Three Performance Measures of Daily Return Data of DOW 30 Stocks.

Name	AS	FH	Sharpe
MMM	0.032	0.032	0.019
AXP	0.013	0.013	0.012
AAPL	0.054	0.049	0.054
BA	0.039	0.038	0.031
CAT	0.021	0.021	0.018
CVX	0.028	0.028	0.020
CSCO	0.019	0.019	0.014
KO	0.075	0.072	0.036
DIS	0.039	0.039	0.028
XOM	0.018	0.018	0.010
GS	0.010	0.010	0.009
HD	0.043	0.042	0.030
IBM	0.020	0.020	0.010
INTC	0.023	0.023	0.018
JNJ	0.065	0.063	0.028
JPM	0.016	0.016	0.017
MCD	0.086	0.083	0.046
MRK	0.036	0.035	0.025
MSFT	0.042	0.041	0.031
NKE	0.056	0.055	0.044
PFE	0.028	0.028	0.015
PG	0.057	0.057	0.026
TRV	0.032	0.031	0.024
UTX	0.037	0.037	0.023
UNH	0.029	0.029	0.026
VZ	0.040	0.040	0.022
V	0.053	0.051	0.048
WMT	0.041	0.041	0.021
WBA	0.013	0.012	0.007
mean	0.037	0.036	0.025

Table 8 presents summary statistics of monthly return data for the DOW 30 stocks. Mean in the last row in Table 8 denotes mean of summary statistics over 29 stocks. Mean of summary statistics ranges from 0.335 in Walgreens to 2.297 in Apple. Mean ranges from the minimum in Walgreens to the maximum in Apple in monthly data, which is the same as in daily data. In the table, mean* denotes the mean derived from the formula where mean in monthly data should be close to $30 \div 7 \times 5$ times mean in daily data if daily returns follow identical distributions and s.d.* denotes the standard deviation derived from the formula where standard deviation in monthly data should be close to $\sqrt{30 \div 7 \times 5}$ times standard deviation in daily data if daily returns follow independent and identical distributions. Standard deviation ranges from 3.982 in Johnson & Johnson to 9.314 in Apple. Standard deviation in Johnson & Johnson is also the minimum in daily data.

Skewness is all negative in monthly data except American Express. The negative skewness of the distribution shows that we may expect frequent small gains and a few large losses. There are only four companies, 3M, American Express, Cisco Systems, and Walgreens, where skewness is larger in monthly data than daily data. Skewness is more negative in the rest of the four companies in monthly data than daily data. Kurtosis shows

DOW components have tails closer to the normal distribution except American Express in monthly data than daily data.

Table 9 presents percentiles of monthly return data for DOW 30 stocks. Mean in the last row in Table 9 denotes mean of percentiles over 29 stocks. The maximum ranges from 10.364 in Johnson & Johnson to 62.866 in American Express. The minimum ranges from −43.479 in Caterpillar to −10.881 in McDonald's. The range of returns widens in monthly data than in daily data. Consequently, most of summary statistics become larger in absolute value in monthly data than in daily data except kurtosis and 40% percentile. This applies to the maximum and minimum. Hence, the minimum shows more severe disastrous observations in monthly data than daily data. Since the AS and FH performance measures are sensitive to losses but insensitive to gains as we saw in the numerical examples in the previous section, negative values of the observation have disproportionately larger adverse effects on the two performance measures than positive values when the former and latter are equal in absolute value.

Table 8. Summary Statistics of Monthly Return Data of DOW 30 Stocks.

Name	Mean	s.d.	Skew	Kurt	Mean*	s.d.*
MMM	0.627	5.646	−0.580	3.720	0.651	6.306
AXP	0.675	8.670	1.261	18.552	0.661	9.945
AAPL	2.297	9.314	−1.052	6.023	2.445	9.427
BA	1.242	7.607	−0.612	4.237	1.258	8.062
CAT	0.901	9.272	−0.826	6.977	0.884	9.129
CVX	0.738	5.693	−0.275	3.080	0.788	7.445
CSCO	0.681	7.349	−0.322	3.748	0.653	8.226
KO	0.920	4.246	−0.512	4.818	0.935	5.010
DIS	1.018	6.163	−0.216	4.410	1.051	7.347
XOM	0.402	5.185	−0.072	4.112	0.425	6.828
GS	0.524	8.598	−0.389	3.874	0.549	10.293
HD	1.132	6.077	−0.403	3.499	1.126	7.286
IBM	0.412	5.954	−0.939	6.227	0.393	6.239
INTC	0.782	6.963	−0.459	3.592	0.779	8.261
JNJ	0.693	3.982	−0.489	4.147	0.720	4.706
JPM	0.952	7.855	−0.589	4.475	0.945	10.877
MCD	1.254	4.270	−0.229	3.192	1.285	5.481
MRK	0.974	5.970	−0.576	4.592	0.928	7.162
MSFT	1.180	6.498	−0.202	3.746	1.190	7.543
NKE	1.660	6.319	−0.373	3.524	1.662	7.715
PFE	0.600	5.325	−0.402	3.595	0.558	6.318
PG	0.713	4.245	−0.362	3.290	0.707	4.951
TRV	0.926	5.087	−0.114	4.049	0.945	7.643
UTX	0.801	5.480	−0.497	3.255	0.801	6.584
UNH	1.144	7.452	−1.449	8.249	1.178	8.985
VZ	0.760	4.926	−0.128	2.809	0.705	5.912
V	1.825	6.214	−0.148	6.859	1.931	8.507
WMT	0.646	4.794	−0.382	4.440	0.647	5.586
WBA	0.335	7.499	−0.093	3.571	0.371	7.700
mean	0.925	6.298	−0.394	4.850	0.937	7.430

In the table, s.d., skew, and kurt denote respectively standard deviation, skewness, and kurtosis. Mean* denotes the mean derived from the formula where mean in monthly data should be close to $30 \div 7 \times 5$ times mean in daily data if daily returns follow identical distributions and s.d.* denotes the standard deviation derived from the formula where standard deviation in monthly data should be close to $\sqrt{30 \div 7 \times 5}$ times standard deviation in daily data if daily returns follow independent and identical distributions.

Table 9. Percentiles of Monthly Return Data of DOW 30 Stocks.

Name	Max	Min	80%	60%	40%	20%
MMM	15.989	−16.231	4.913	2.408	0.253	−3.516
AXP	62.866	−32.731	5.340	2.159	−0.457	−3.182
AAPL	21.326	−39.982	9.476	5.449	0.285	−4.360
BA	20.383	−28.691	7.151	3.344	0.021	−4.793
CAT	30.032	−43.479	7.737	3.206	−0.447	−5.404
CVX	14.161	−15.897	5.210	2.604	−0.608	−3.659
CSCO	20.691	−23.867	6.193	2.395	−0.732	−4.668
KO	13.272	−18.247	4.355	2.016	0.025	−2.538
DIS	20.996	−20.958	5.587	2.324	−0.135	−3.396
XOM	20.941	−15.346	4.644	1.531	−0.618	−3.781
GS	21.036	−32.178	7.259	2.606	−0.824	−6.771
HD	14.889	−18.053	6.214	2.577	−0.136	−3.801
IBM	16.767	−27.001	5.240	1.707	−0.526	−3.697
INTC	17.780	−23.389	5.883	3.353	−0.287	−4.953
JNJ	10.364	−13.480	3.470	1.670	−0.380	−2.436
JPM	21.813	−26.460	6.760	3.223	−0.520	−4.914
MCD	13.037	−10.881	4.664	2.494	0.541	−1.588
MRK	15.761	−23.153	5.515	2.504	−0.031	−3.064
MSFT	22.274	−17.836	6.109	2.872	0.492	−4.047
NKE	17.430	−20.896	6.782	3.369	0.062	−3.287
PFE	13.964	−19.448	5.115	1.955	−0.851	−3.574
PG	10.981	−12.346	4.382	1.830	−0.192	−2.511
TRV	18.017	−15.681	4.722	2.086	0.232	−3.572
UTX	12.782	−15.311	5.042	2.617	−0.154	−2.988
UNH	23.591	−36.584	6.589	3.323	0.946	−4.102
VZ	13.114	−12.451	5.172	1.754	−0.712	−3.098
V	29.132	−21.784	5.892	3.609	0.989	−2.489
WMT	13.782	−17.373	4.211	1.470	−0.469	−2.955
WBA	22.016	−21.441	6.138	2.234	−1.612	−5.372
mean	19.627	−22.110	5.716	2.576	−0.202	−3.742

In the table, max, min, 80%, 60%, 40%, and 20% denote respectively maximum, minimum, 80% percentile, 60% percentile, 40% percentile, and 20% percentile.

Table 10 presents the three performance measures of monthly return data for the DOW 30 stocks. The Sharpe ratio is high in monthly data in some stocks such as Visa, McDonald's, Nike, and Apple. The Sharpe ratio in monthly data is much higher than that in daily data. This is natural since the monthly Sharpe ratio is close to $\sqrt{30 \div 7 \times 5}$ times as much as the daily Sharpe ratio if daily returns are independently and identically distributed (cf. Lo 2002), where 30 denotes an average number of days in a month, 7 denotes the number of days in a week, and 5 denotes the number of weekdays in a week. On the other hand, the AS and FH performance measures in monthly data are much closer to those in daily data. They are nearly closed under temporal aggregation in some stocks, i.e., they have time-invariant values regardless of data frequency (cf. Hodoshima 2020b).

The AS performance measure ranges from 0.012 in Walgreens to 0.132 in McDonald's. McDonald's is the maximum in the AS performance measure in monthly data, which is the same as in daily data. McDonald's is by far the best by the AS performance measure but rated only as the second-best next to Visa by the Sharpe ratio. The difference between the

AS performance measure and Sharpe ratio is large in outperforming stocks but small in underperforming stocks.

Table 10. Three Performance Measures of Monthly Return Data of DOW 30 Stocks.

Name	AS	FH	Sharpe	Sharpe*	dailyAS	dailyFH
MMM	0.038	0.035	0.093	0.086	0.032	0.032
AXP	0.019	0.018	0.066	0.056	0.013	0.013
AAPL	0.044	0.025	0.235	0.248	0.054	0.049
BA	0.040	0.032	0.149	0.143	0.039	0.038
CAT	0.020	0.017	0.086	0.085	0.021	0.021
CVX	0.045	0.042	0.111	0.091	0.028	0.028
CSCO	0.025	0.024	0.078	0.066	0.019	0.019
KO	0.094	0.055	0.192	0.165	0.075	0.072
DIS	0.052	0.043	0.148	0.128	0.039	0.039
XOM	0.030	0.029	0.057	0.047	0.018	0.018
GS	0.014	0.014	0.049	0.043	0.010	0.010
HD	0.059	0.048	0.169	0.140	0.043	0.042
IBM	0.022	0.021	0.052	0.046	0.020	0.020
INTC	0.031	0.029	0.097	0.081	0.023	0.023
JNJ	0.083	0.066	0.148	0.130	0.065	0.063
JPM	0.030	0.026	0.108	0.077	0.016	0.0160
MCD	0.132	0.092	0.269	0.215	0.086	0.083
MRK	0.051	0.040	0.145	0.115	0.036	0.035
MSFT	0.055	0.047	0.165	0.143	0.042	0.041
NKE	0.078	0.048	0.246	0.201	0.056	0.055
PFE	0.041	0.038	0.093	0.071	0.028	0.028
PG	0.076	0.066	0.143	0.121	0.057	0.057
TRV	0.070	0.058	0.161	0.110	0.032	0.031
UTX	0.051	0.046	0.127	0.105	0.037	0.037
UNH	0.036	0.026	0.139	0.119	0.029	0.029
VZ	0.062	0.057	0.133	0.101	0.040	0.040
V	0.086	0.046	0.287	0.222	0.053	0.051
WMT	0.054	0.047	0.113	0.096	0.041	0.041
WBA	0.012	0.012	0.031	0.034	0.013	0.012
mean	0.050	0.039	0.134	0.113	0.037	0.036

In the table, Sharpe* denotes the Sharpe ratio under the formula where the Sharpe ratio in monthly data should be close to $\sqrt{30} \div 7 \times 5$ times the Sharpe ratio in daily data if daily returns follow independent and identical distributions. Also in the table, dailyAS and daily FH denote the AS performance measure and FH performance measure in daily data respectively to compare.

The FH performance measure is small compared to the AS performance measure. Some companies with high AS performance measure scores in monthly data referred above all have the FH performance measure considerably smaller than the corresponding AS performance measure. This is in contrast to the result in daily data. As we saw in the previous subsection, the FH performance measure is much more sensitive to losses of the underlying stock performance, we consider this indicates the lower FH performance measure score in monthly data is due to larger losses in the stock return in monthly data. Therefore, our result indicates the FH performance measure is more sensitive to losses, i.e., negative skewness or left tail of the underlying distribution, than the AS performance measure.

4. Discussion

In this section, in order to see how the AS and FH performance measures are related to summary statistics of the underlying return data, we run regressions with each of the two performance measures as the dependent variable and summary statistics as explanatory variables in both daily and monthly data. We add percentiles of the maximum, minimum, 80%, 60%, 40%, and 20% to summary statistics in regressions to see how these percentiles are related to the two performance measures. This is to examine how the two performance measures are sensitive to losses or maximal loss of the underlying return data. In particular, in this section we focus on the FH performance measure more than the AS performance measure since the former is known to be extremely sensitive to disaster risk or maximal loss as compared to the latter.

We begin to compare the estimate of the FH performance measure and the reciprocal of the maximal loss since the maximal loss is identified as the extended FH risk measure when there exists no FH risk measure (cf. [Riedel and Hellmann 2015](#)).

Table 11 shows, in daily data, the FH performance measure estimate, the reciprocal of the maximal loss L , and the sample estimate of the expectation $E[\ln(1 + \frac{\xi}{L})]$, which becomes a discriminant if the reciprocal of the maximal loss should be used as the extended FH performance measure or not. Proposition 3 in Section 2 implies the FH risk (performance) exists when $E[\ln(1 + \frac{\xi}{L})]$ becomes negative but does not exist when $E[\ln(1 + \frac{\xi}{L})]$ becomes nonnegative. If the discriminant $E[\ln(1 + \frac{\xi}{L})]$ becomes nonnegative, then (the reciprocal of) the maximal loss should be used as the extended FH risk (performance) measure. The estimate of the discriminant becomes positive only once in Apple and is negative otherwise. Therefore, the GMM estimate of the FH performance measure should be used as the FH performance measure except for Apple. In every stock in Table 11, the GMM estimate of the FH performance measure is uniformly smaller than the estimate of the extended FH performance measure, i.e., the reciprocal of the maximal loss. The difference between the two estimates is in general quite large so that one should use the GMM estimate as the FH performance measure. Even in Apple's case where one should use the reciprocal of the maximal loss, the difference between the two estimates is very small so that the use of the GMM estimate as the FH performance measure estimate in this case is not problematic.

Table 12 shows, in monthly data, the FH performance measure estimate, the reciprocal of the maximal loss, and the sample estimate of the expectation $E[\ln(1 + \frac{\xi}{L})]$. The discriminant estimate takes positive values in 12 stocks where one should use the reciprocal of the maximal loss as the extended FH performance measure. However, the difference between the GMM estimate of the FH performance measure and the reciprocal of the maximal loss is generally small in these cases but large when the discriminant estimate takes negative values. Therefore, the erroneous use of the reciprocal of the maximal loss as the extended FH performance measure results in bad inference of the FH performance measure when the discriminant estimate takes negative values while the incorrect use of the GMM estimate of the FH performance measure does not cause much troubles when the discriminant estimate takes positive values. This view is new in the existing literature. We emphasize the importance of finding the GMM estimate by grid search when we estimate the FH performance (risk) measure in order to obtain a good estimate.

We now turn to regression analysis in order to find how sensitive the new performance measures are to disaster risk or maximal loss. Table 13 shows the regression result of the AS performance measure and summary statistics in daily return data. We test whether each regression coefficient is zero or not by a t -test in regression analysis. We name explanatory variables of percentiles as risks such as risk0, risk1, etc. in the tables of regression results given below. The goodness of fit statistics, shown by R^2 being 0.907 and \bar{R}^2 being 0.855, indicate this regression fits data fairly well. Mean is one percent significant and has the largest estimate of 0.503. Significance and insignificance in the regression results for the regression of the two performance measures and summary statistics in this section are based on p -values under the classical assumptions of the standard regression model of

homoskedastic variance of the error term and independently and identically normally distributed error term so that t -values are assumed to follow t distribution with 18 degrees of freedom. This is so because we are concerned about the finite-sample inference of the regression with only 29 observations and 10 explanatory variables, where there are not many observations given the number of explanatory variables in the regression model. Standard deviation is five percent significant with the second-largest estimate of -0.043 in absolute value. Skewness has a ten percent significant positive estimate of 0.023 but kurtosis is not significant. Other explanatory variables are not significant. Six percentiles have all small t -values in absolute values, and hence they are not significant. Therefore, only mean, standard deviation, and skewness are significant.

Table 11. The FH Performance Measure Estimate, Reciprocal of the Maximal Loss, and Estimate of $E[\ln(1 + \frac{\xi}{L})]$ for Daily Return Data of the DOW 30 Stocks.

Name	FH	Reciprocal of Maximal Loss	Estimate of the Discriminant
MMM	0.032	0.072	−0.0026
AXP	0.013	0.052	−0.0048
AAPL	0.049	0.051	0.0007
BA	0.038	0.107	−0.0131
CAT	0.021	0.064	−0.0055
CVX	0.028	0.075	−0.0057
CSCO	0.019	0.057	−0.0034
KO	0.072	0.110	−0.0027
DIS	0.039	0.098	−0.0091
XOM	0.018	0.067	−0.0034
GS	0.010	0.048	−0.0044
HD	0.042	0.117	−0.0119
IBM	0.020	0.115	−0.0128
INTC	0.023	0.076	−0.007
JNJ	0.063	0.095	−0.0012
JPM	0.016	0.043	−0.0032
MCD	0.083	0.120	−0.003
MRK	0.035	0.063	−0.0019
MSFT	0.041	0.080	−0.0052
NKE	0.055	0.079	−0.0027
PFE	0.028	0.089	−0.0057
PG	0.057	0.122	−0.0048
TRV	0.031	0.050	−0.0009
UTX	0.037	0.109	−0.0088
UNH	0.029	0.048	−0.0016
VZ	0.040	0.119	−0.0088
V	0.051	0.068	−0.0019
WMT	0.041	0.093	−0.004
WBA	0.012	0.062	−0.0046

Table 11 shows the FH performance measure estimate, the reciprocal of the maximal loss, and the estimate of the discriminant $E[\ln(1 + \frac{\xi}{L})]$ for daily return data of the DOW 30 stocks. The estimate of the discriminant $E[\ln(1 + \frac{\xi}{L})]$ is obtained from the sample data where the observation where the maximal loss occurs is excluded from the sample not to make the estimate take $-\infty$.

Table 12. The FH Performance Measure Estimate, Reciprocal of the Maximal Loss, and Estimate of $E[\ln(1 + \frac{x}{L})]$ for Monthly Return Data of the DOW 30 Stocks.

Name	FH	Reciprocal of Maximal Loss	Estimate of the Discriminant
MMM	0.035	0.062	−0.0464
AXP	0.018	0.031	−0.0081
AAPL	0.025	0.025	0.0292
BA	0.032	0.035	0.0118
CAT	0.017	0.023	0.0029
CVX	0.042	0.063	−0.0224
CSCO	0.024	0.042	−0.0202
KO	0.055	0.055	0.0314
DIS	0.043	0.048	0.0087
XOM	0.029	0.065	−0.0338
GS	0.014	0.031	−0.0151
HD	0.048	0.055	−0.0022
IBM	0.021	0.037	−0.0058
INTC	0.029	0.043	−0.0104
JNJ	0.066	0.074	0.0007
JPM	0.026	0.038	−0.0222
MCD	0.092	0.092	0.0341
MRK	0.040	0.043	0.0135
MSFT	0.047	0.056	−0.0212
NKE	0.048	0.048	0.0388
PFE	0.038	0.051	−0.0019
PG	0.066	0.081	−0.0119
TRV	0.058	0.064	0.0086
UTX	0.046	0.065	−0.0250
UNH	0.026	0.027	0.0138
VZ	0.057	0.080	−0.0254
V	0.046	0.046	0.0405
WMT	0.047	0.058	−0.0052
WBA	0.012	0.047	−0.0546

Table 12 shows the FH performance measure estimate, the reciprocal of the maximal loss, and the estimate of the discriminant $E[\ln(1 + \frac{x}{L})]$ for monthly return data of the DOW 30 stocks. The estimate of the discriminant $E[\ln(1 + \frac{x}{L})]$ is obtained from the sample data where the observation where the maximal loss occurs is excluded from the sample not to make the estimate take $-\infty$.

Table 13. Regression Result of the AS Performance Measure and Summary Statistics for Daily Return Data of the DOW 30 Stocks.

	Intcept	Mean	s.d.	Skew	Kurt	Risk0	Risk5	Risk1
estimate	0.112	0.503	−0.043	0.023	−0.002	0.001	−0.002	−0.019
t-value	5.858 ***	3.034 ***	−2.208 **	1.847 *	−1.248	0.450	−1.359	−0.224
	Risk2	Risk3	Risk4	R^2	\bar{R}^2			
estimate	0.026	0.091	0.012	0.907	0.855			
t-value	0.254	0.887	0.166					

Table 13 shows the regression result of the AS performance measure as the dependent variable and summary statistics as explanatory variables for daily return data of the DOW 30 stocks. In the table, s.d., skew, kurt, risk0, and risk5 denote respectively standard deviation, skewness, kurtosis, maximum, and minimum. Also risk1, risk2, risk3, and risk4 denote respectively 80% percentile, 60% percentile, 40% percentile, and 20% percentile. We name percentiles of the return distribution as classes of risk markers such as risk5 etc. In the table, R^2 and \bar{R}^2 are goodness of fit statistics, i.e., R-square and adjusted R-square respectively. In the table, *** denotes significant at 1% level, ** denotes significant at 5% level, and * denotes significant at 10% level. Significance and insignificance in the regression results for the regression of the two performance measures and summary statistics are based on p -values under the classical assumptions of the standard regression model of homoskedastic variance of the error term and independently and identically normally distributed error term so that t -values are assumed to follow t distribution with 18 degrees of freedom.

Table 14 shows the regression result of the FH performance measure and summary statistics in daily data. The goodness of fit in the regression, given by R^2 and \bar{R}^2 , is fairly good, similar to the result for the AS performance measure. The estimation result in Table 14 is similar to that in Table 13. Mean is highly significant with the largest estimate of 0.470. Standard deviation is ten percent significant with the second-largest t-value in absolute value. Skewness is also ten percent significant and has a positive estimate, while kurtosis is not significant. None of six percentiles are significant in Table 14, which is the same as in Table 13. Therefore, mean, standard deviation, and skewness are the most important factors influencing the FH performance measure in daily data. However, percentiles, including the maximum and minimum, are not significant in daily data despite the previous view the FH performance measure is extremely sensitive to rare disasters and our numerical investigation in Section 3. Overall, the regression result of the FH performance measure in Table 14 is similar to that in the AS performance measure in Table 13 for daily data.

Table 14. Regression Result of the FH Performance Measure and Summary Statistics for Daily Return Data of the DOW 30 Stocks.

	Intcept	Mean	s.d.	Skew	Kurt	Risk0	Risk5	Risk1
estimate	0.109	0.470	−0.040	0.022	−0.002	0.001	−0.002	−0.028
t-value	5.882 ***	2.907 ***	−2.074 *	1.798 *	−1.208	0.449	−1.212	−0.328
	Risk2	Risk3	Risk4	R^2	\bar{R}^2			
estimate	0.025	0.089	0.003	0.904	0.850			
t-value	0.251	0.895	0.038					

Table 14 shows the regression result of the FH performance measure as the dependent variable and summary statistics as explanatory variables for daily return data of the DOW 30 stocks. In the table, s.d., skew, kurt, risk0, and risk5 denote respectively standard deviation, skewness, kurtosis, maximum, and minimum. Also risk1, risk2, risk3, and risk4 denote respectively 80% percentile, 60% percentile, 40% percentile, and 20% percentile. We name percentiles of the return distribution as classes of risk markers such as risk5 etc. In the table, R^2 and \bar{R}^2 are goodness of fit statistics, i.e., R-square and adjusted R-square respectively. In the table, *** denotes significant at 1% level, and * denotes significant at 10% level. Significance and insignificance in the regression results for the regression of the two performance measures and summary statistics are based on p -values under the classical assumptions of the standard regression model of homoskedastic variance of the error term and independently and identically normally distributed error term so that t -values are assumed to follow t distribution with 18 degrees of freedom.

Table 15 shows the regression result of the AS performance measure and summary statistics in monthly data. The goodness of fit statistics are a little smaller in monthly data than in daily data. Mean and standard deviation, the most influential factors in daily data, are no longer significant in monthly data. The only significant factor, significant at ten percent, is skewness with p -value 0.055 in monthly data. The positive estimate of skewness implies skewness has a negative effect on the AS performance measure since the majority of the 29 companies have negative skewness in monthly data, as we saw above. The estimate 0.045 of skewness is the largest estimate in Table 15. We consider the negative effect of skewness in Table 15 indicates the AS performance measure is sensitive to losses or the underlying risk of stock returns (cf., e.g., Kadan and Liu 2014; Miyahara 2014; Ban et al. 2016; Hodoshima 2019). As the reason why mean and standard deviation are no longer significant in monthly data, we consider daily data contain daily microstructure noise and can be summarized by traditional summary statistics such as mean and standard deviation but that monthly data, daily microstructure noise being washed out, have distributions where only skewness is significant when explanatory variables of percentiles, including the maximum and minimum, are present.

Table 16 shows the regression result of the FH performance measure and summary statistics in monthly data. The goodness of fit statistics are also a little smaller in monthly data than in daily data. Mean and standard deviation, the two most influential factors in daily data, are no longer significant in monthly data as in the regression of the AS performance measure in Table 15. Monthly return data have distributions where traditional summary statistics of mean and standard deviation lose explanatory power in

the regression when other percentile summary statistics are present. Instead, skewness and 40% percentile are five percent significant with p -value 0.023 and 0.030 respectively. The maximum is ten percent significant. However, the maximum has a small estimate -0.002 so that it is not an important factor. Skewness has the largest estimate of 0.036, and hence it is the most influential factor. Since the majority of the 29 companies have negative skewness in monthly data, the positive estimate of skewness implies skewness has a negative effect on the FH performance measure. The 40% percentile being five percent significant seems to indicate the FH performance measure is more sensitive to losses of stock returns compared to the AS performance measure in monthly data. However, the minimum has a small estimate and a low t -value. This implies the FH performance measure is not sensitive to the maximal loss or the rare disaster, which has been used as the value of the FH extended riskiness or performance measure in some previous studies (cf., Kadan and Liu 2014; Anand et al. 2016, 2017; Riedel and Hellmann 2015). This is evidence that the FH performance is more sensitive to losses of the underlying financial target compared to the AS performance measure but not excessively determined by the maximal loss. Unlike the previous studies in the literature and our own numerical investigation in Section 3, our empirical results show the FH performance measure is not excessively determined by the maximal loss.

Table 15. Regression Result of the AS Performance Measure and Summary Statistics for Monthly Return Data of the DOW 30 Stocks.

	Intcept	Mean	s.d.	Skew	Kurt	Risk0	Risk5	Risk1
estimate	0.147	0.028	−0.006	0.045	−0.003	−0.002	−0.003	−0.006
t -value	6.357 ***	1.005	−0.432	2.050 *	−0.744	−1.023	−1.716	−0.469
	Risk2	Risk3	Risk4	R^2	\bar{R}^2			
	−0.002	0.016	0.009	0.882	0.816			
	−0.209	1.523	1.104					

Table 15 shows the regression result of the AS performance measure as the dependent variable and summary statistics as explanatory variables for monthly return data of the DOW 30 stocks. In the table, s.d., skew, kurt, risk0, and risk5 denote respectively standard deviation, skewness, kurtosis, maximum, and minimum. Also risk1, risk2, risk3, and risk4 denote respectively 80% percentile, 60% percentile, 40% percentile, and 20% percentile. We name percentiles of the return distribution as classes of risk markers such as risk5 etc. In the table, R^2 and \bar{R}^2 are goodness of fit statistics, i.e., R-square and adjusted R-square respectively. In the table, *** denotes significant at 1% level, and * denotes significant at 10% level. Significance and insignificance in the regression results for the regression of the two performance measures and summary statistics are based on p -values under the classical assumptions of the standard regression model of homoskedastic variance of the error term and independently and identically normally distributed error term so that t -values are assumed to follow t distribution with 18 degrees of freedom.

Table 16. Regression Result of the FH Performance Measure and Summary Statistics for Monthly Return Data of the DOW 30 Stocks.

	Intcept	Mean	s.d.	Skew	Kurt	Risk0	Risk5	Risk1
estimate	0.109	0.005	0.004	0.036	−0.000	−0.002	−0.001	−0.002
t -value	7.054 ***	0.285	0.439	2.475 ***	−0.127	−2.049 *	−0.973	−0.221
	Risk2	Risk3	Risk4	R^2	\bar{R}^2			
	−0.005	0.016	0.009	0.883	0.818			
	−0.707	2.364 ***	1.639					

Table 16 shows the regression result of the FH performance measure as the dependent variable and summary statistics as explanatory variables for monthly return data of the DOW 30 stocks. In the table, s.d., skew, kurt, risk0, and risk5 denote respectively standard deviation, skewness, kurtosis, maximum, and minimum. Also risk1, risk2, risk3, and risk4 denote respectively 80% percentile, 60% percentile, 40% percentile, and 20% percentile. We name percentiles of the return distribution as classes of risk markers such as risk5 etc. In the table, R^2 and \bar{R}^2 are goodness of fit statistics, i.e., R-square and adjusted R-square respectively. In the table, *** denotes significant at 1% level, and * denotes significant at 10% level. Significance and insignificance in the regression results for the regression of the two performance measures and summary statistics are based on p -values under the classical assumptions of the standard regression model of homoskedastic variance of the error term and independently and identically normally distributed error term so that t -values are assumed to follow t distribution with 18 degrees of freedom.

5. Concluding Comments

We examined how sensitive the new performance measures of the AS and FH performance measures are to *disaster risk* or *maximal loss*. We presented numerical examples and empirical examples. Although numerical examples are limited, we showed that the AS performance measure is sensitive to *disaster risk* but insensitive to gains. On the other hand, we showed that the FH performance measure is even more sensitive to *disaster risk* than the AS performance measure and more sensitive to gains than the AS performance measure. The FH performance measure is sometimes not possible to obtain by the large maximal loss. On the other hand, our empirical examples show we can always obtain the two performance measures despite the large maximal loss. Therefore, the infeasibility of the FH performance measure due to the large maximal loss does not happen in our representative stock return data.

We closely studied how the GMM estimate of the FH performance measure is related to the maximal loss, the reciprocal of which is the extended FH performance measure defined by [Riedel and Hellmann \(2015\)](#) when the FH risk (performance) measure does not exist. We found that the GMM estimate is valid in every DOW 30 stocks except Apple in daily DOW 30 stock data. The difference between the GMM estimate and the reciprocal of the maximal loss (the extended FH performance measure) in daily data is very small in Apple but large in other stocks. Therefore, we should be careful to use the extended FH risk (performance) measure not to make erroneous estimation. On the other hand, the GMM estimate of the FH performance measure is more often invalid in monthly DOW 30 stock data. However, the difference between the GMM estimate and the reciprocal of the maximal loss in monthly data is small when the FH risk (performance) measure does not exist but large when the FH risk (performance) measure exists and hence (the GMM estimate of) the FH performance measure is valid. Therefore, one should check the sign of the discriminant $E\left[\ln\left(1 + \frac{g}{L}\right)\right]$ whether one should use the FH performance measure or the extended FH performance measure, i.e., the reciprocal of the maximal loss. However, (the GMM estimate of) the FH performance measure does not function so badly when it is not valid, i.e., when the FH risk (performance) measure does not exist, and does much better than the extended FH performance measure when it is valid. Therefore, we do not make a big mistake by (the GMM estimate of) the FH performance measure.

We also examined whether the two performance measures are related to summary statistics including percentiles by running regressions with the two performance measures as dependent variables and summary statistics as explanatory variables. In daily data, only mean and standard deviation are respectively one and five percent significant for the AS performance measure and only mean is one percent significant for the FH performance measure. In monthly data, only skewness is ten percent significant for the AS performance measure and skewness and 40% percentile are five percent significant for the FH performance measure. Therefore, the maximal loss or disaster risk is not, unlike the previous view in the literature, related to the two performance measures. However, skewness and 40% percentile being five percent significant in the FH performance measure as compared to skewness being only 10 percent significant in the AS performance measure indicates the FH performance measure is more sensitive to losses in return observations compared to the AS performance measure. Therefore, the two performance measures can provide evaluations in representative stock data, even though they contain quite disastrous observations, despite the previous view in the literature and our numerical examples. Hence, the two performance measures could be used in empirical studies to shed new light to show risk-averse assessments as compared to other traditional performance measures.

In this study, we identified risk as losses. However, risk is sometimes associated with regional economic concepts such as regional risks for doing business. For example, the World Economic Forum has been publishing The Global Risks Report (cf., e.g., [The Global Risks Report \(2020\)](#)) since 2006, highlighting each year vulnerability of our world to volatility and disruption. Studies such as ours where risk is associated with losses may be limited when we consider risk more broadly.

Author Contributions: Conceptualization, J.H. and T.Y.; Methodology, J.H. and T.Y.; Software, J.H. and T.Y.; Validation, J.H. and T.Y.; Formal Analysis, J.H.; Investigation, J.H. and T.Y.; Resources, J.H. and T.Y.; Data Curation, T.Y.; Writing—Original Draft Preparation, J.H.; Writing—Review & Editing, J.H. and T.Y.; Visualization, J.H. and T.Y.; Supervision, J.H.; Project Administration, J.H.; Funding Acquisition, J.H. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by JSPS KAKENHI Grant Number JP17K03667.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Anand, Abhinav, Tiantian Li, Tetsuo Kurosaki, and Young-Shin Kim. 2016. Foster-Hart Optimal Portfolios. *Journal of Banking and Finance* 68: 117–30. [\[CrossRef\]](#)
- Anand, Abhinav, Tiantian Li, Tetsuo Kurosaki, and Young-Shin Kim. 2017. The equity risk by the too-big-to-fail banks: A Foster-Hart estimation. *Annals of Operations Research* 253: 21–41. [\[CrossRef\]](#)
- Aumann, Robert J., and Roberto Serrano. 2008. An economic index of riskiness. *Journal of Political Economy* 116: 810–36. [\[CrossRef\]](#)
- Ban, Lan, Tetsuya Misawa, and Yoshio Miyahara. 2016. Valuation of Hong Kong REIT based on risk sensitive value measure method. *International Journal of Real Options and Strategy* 4: 1–33. [\[CrossRef\]](#)
- Cogneau, Philippe, and Georges Hubner. 2009a. The (more than) 100 ways to measure portfolio performance. Part 1: Standardized risk-adjusted measure. *Journal of Performance Measurement* 13: 56–71.
- Cogneau, Philippe, and Georges Hubner. 2009b. The (more than) 100 ways to measure portfolio performance. Part 2: Special measures and comparison. *Journal of Performance Measurement* 14: 56–69.
- Eling, Martin, and Frank Schuhmacher. 2007. Does the choice of performance measure influence the evaluation of hedge funds? *Journal of Banking and Finance* 31: 2632–47. [\[CrossRef\]](#)
- Farinelli, Simone, Manuel Ferreira, Damiano Rossello, Markus Thoeny, and Luisa Tibiletti. 2008. Beyond Sharpe ratio: Optimal asset allocation using different performance ratios. *Journal of Banking and Finance* 32: 2057–63. [\[CrossRef\]](#)
- Foster, Dean P., and Sergiu Hart. 2009. An operational measure of riskiness. *Journal of Political Economy* 117: 785–814. [\[CrossRef\]](#)
- Hart, Sergiu. 2011. Comparing Risks by Acceptance and Rejection. *Journal of Political Economy* 119: 617–38 [\[CrossRef\]](#)
- Hodoshima, Jiro. 2019. Stock performance by utility indifference pricing and the Sharpe ratio. *Quantitative Finance* 19: 327–38. [\[CrossRef\]](#)
- Hodoshima, Jiro. 2020a. Evaluation of performance of stock and real estate investment trust markets in Japan. *Empirical Economics*. [\[CrossRef\]](#)
- Hodoshima, Jiro. 2020b. Temporal aggregation of the Aumann-Serrano and Foster-Hart performance indexes. Nagoya University of Commerce and Business Discussion Paper. Available online: <https://www.nucba.ac.jp/university/library/discussion-paper/NUCB-DP-20001.html> (accessed on 10 November 2020).
- Hodoshima, Jiro, and Yoshio Miyahara. 2020. Utility indifference pricing and the Aumann-Serrano performance index. *Journal of Mathematical Economics* 86: 83–89. [\[CrossRef\]](#)
- Hodoshima, Jiro, and Toshiyuki Yamawake. 2020. Winners and losers in DOW 30 stocks by the performance indexes incorporating high moments and disaster risk. *Applied Economics*. under review.
- Kadan, Ohad, and Fang Liu. 2014. Performance evaluation with high moments and disaster risk. *Journal of Financial Economics* 113: 131–55. [\[CrossRef\]](#)
- Lo, Andrew W. 2002. The statistics of Sharpe ratios. *Financial Analysts Journal* 58: 36–52. [\[CrossRef\]](#)
- Miyahara, Yoshio. 2010. Risk-sensitive value measure method for projects evaluation. *Journal of Real Options and Strategy* 3: 185–204. [\[CrossRef\]](#)
- Miyahara, Yoshio. 2014. Evaluation of the scale risk. *RIMS Kokyuroku* 1886: 181–88.
- Riedel, Frank, and Tobias Hellmann. 2015. The Foster-Hart measure of riskiness for general gambles. *Theoretical Economics* 10: 1–9. [\[CrossRef\]](#)
- The Global Risks Report. 2020. Available online: <https://www.weforum.org/reports/the-global-risks-report-2020> (accessed on 1 February 2021).