Optimal Hedge Ratio under a Subjective Re-weighting of the Original Measure

MASSIMILIANO BARBI and SILVIA ROMAGNOLI*

THIS VERSION: February 8, 2012

Abstract

In this paper we propose a risk-minimizing optimal hedge ratio with futures contracts, where the risk of hedged portfolio is measured by a fully coherent distortion risk measure, i.e. a parametric coherent risk measure (in the usual meaning introduced by Artzner et al., 1997, 1999) satisfying also a second level of coherence related to the parameter that models agent’s risk aversion or pessimism. Using simulated data we show how the optimal hedge ratio is affected by the characteristics of the underlying returns probability distribution and the agent’s risk aversion. Our results confirm that coherently weighting agent’s risk aversion plays a substantial role.

JEL classification: G11; G32.

Keywords: Optimal hedge ratio; Spectral risk measures; Futures hedging.

*The authors are from the University of Bologna (Italy), Department of Management (Barbi) and Department of Statistics (Romagnoli), respectively. Corresponding author: Silvia Romagnoli, Viale Filopanti 5, 40126 Bologna, Italy; e-mail address: silvia.romagnoli@unibo.it. We would like to thank Emanuele Bajo for his helpful comments. All errors are our responsibility.
1 Introduction

Risk management with derivative contracts has gained increasing importance in the last years. Despite the massive use of derivatives in the nowadays financial markets, the most noble function of these instruments remains to protect economic agents from price variability, that is hedging cash positions. Futures trading allows to hedge positions in the spot market in a straightforward way, thanks to the linearity of its payoff. The practical (naive) recommendation to protect a long (short) cash position is to enter in an equal size short (long) position in the futures market. However, such a “rule of thumb” is unsatisfactory from a scientific viewpoint and thus the determination of the “optimal” hedge ratio (OHR)—that is the size of the futures position to be held in order to hedge a unitary spot position—still remains a debated research direction in the financial literature. A huge number of studies has dealt with this problem. As Chen et al. (2003) point out, the determination of the OHR can be reconducted to the choice of the proper objective function to be optimized.

Basically, the literature classifies the OHR as being either risk-minimizing or utility-maximizing. Measuring risk with the variance of the hedged portfolio is the most common choice, since Johnson (1960) and Stein (1961). Thanks to the fact that the minimum-variance hedge ratio (MVHR) is simple to understand and easy to estimate, it is perhaps the most widely-used hedging strategy. However, the MVHR has a number of unattractive theoretical characteristics, mainly the use of variance as a measure of portfolio risk is reasonable only when agents have either quadratic utility preferences or returns are drawn from a multivariate normal distribution, and neither of these two assumptions is justified in practice. In fact, quadratic preferences imply negative marginal utility after a certain level of wealth and increasing absolute risk aversion, and therefore they are not representative of investor’s behavior. Moreover, empirical financial literature strongly rejects the assumption of normally distributed returns, and there is evidence that moments higher than two are priced by the market (Kraus and Litzenberger, 1976; Harvey and Siddique, 2000; Dittmar, 2002). To circumvent this shortcoming, alternative risk measures have been proposed (see Chen et al., 2003, for a review of such approaches). One of the most popular risk measure is value-at-risk (VaR), whose major drawback is that it is not always subadditive, nor convex. Subadditivity, in particular, is the most important property we

---

1 Given a probability space \((\Omega, \mathcal{F}, \mathbb{P})\) and two random variables \(X_1, X_2 \in L_2(\mathbb{P})\), if \(\rho\) is a risk measure, then subadditivity means that \(\rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2)\), while convexity implies that \(\rho(\lambda X_1 + (1 - \lambda) X_2) \leq \lambda \rho(X_1) + (1 - \lambda) \rho(X_2), \lambda \in [0,1]\). As observed by Föllmer and Schied (2002), the notion of convexity synthesizes that of subadditivity and positive homogene-
would expect a risk measure to satisfy as it reflects the common idea of risk diversification. Artzner et al. (1997, 1999) characterize axiomatically the properties of a coherent risk measure. Since subadditivity is required, VaR is not a coherent measure of risk.\(^2\)

Instead of minimizing a measure of portfolio risk, some studies (Cecchetti et al., 1988; Lence, 1995, 1996) propose to derive an OHR based on the maximization of the agent’s expected utility. In this way, the subjective risk aversion of the agent is taken into account and the estimated OHR depends on the chosen agent’s utility function. As proved by Benninga et al. (1983) (see also Lien and Tse, 2002, for a discussion), the maximum-expected utility hedge ratio reduces to the MVHR under joint normality of spot and futures returns.\(^3\) Interestingly, the link between risk-minimizing and utility-maximizing hedge ratio has not been emphasized in the literature. In particular, leaving the trivial case of normal distributed returns aside, to our knowledge the agent’s perception of risk has never been incorporated into a risk-minimizing hedge ratio. In this paper we propose to derive a risk-minimizing hedge ratio employing a new class of risk measures which incorporates agent’s subjective preferences towards risk.

The most general result on the subjective characterization of a risk measure in terms of agent’s risk aversion (or pessimism) is the class of the so-called distortion risk measures (DRMs) introduced by Denneberg (1994) and Wang et al. (1997). A DRM is defined as the expected loss under a transformation of the underlying cumulative density function by means of a given distortion function. Wang et al. (1997) present the properties which characterize a DRM, i.e. monotonicity, positive homogeneity, translational invariance and comonotonic additivity.\(^4\) It can be shown (Yaari, 1987; Schmeidler, 1989) that if and only if a risk measure satisfies comonotonic additivity it has a representation as a Choquet integral. Therefore, DRMs are a special class of the so-called Choquet expected utility, (stating that \(\rho(\lambda X_1) = \lambda \rho(X_1), \lambda \geq 0\)).

\(^2\)An excellent review of the theory of coherent risk measures in risk management applications is Acerbi (2007).

\(^3\)The conditions under which maximum-expected utility hedge ratio is equal to MVHR are the following: (i) futures prices are unbiased (i.e., futures prices follow a martingale), and (ii) spot returns can be written as a linear function of futures returns, and futures returns are independent of residual randomness of spot returns. Condition (i) is generally accepted, and condition (ii) holds in case spot and futures returns are jointly normally distributed.

\(^4\)Formally, given a probability space \((\Omega, \mathcal{F}, \mathbb{P})\) and a random variable \(X_1 \in L_2(\mathbb{P})\), monotonicity requests that \(\rho(X_1) \geq 0\) if \(X_1 \geq 0\), positive homogeneity means that \(\rho(\lambda X_1) = \lambda \rho(X_1), \forall \lambda \geq 0\), while translational invariance means that \(\rho(X_1 + c) = \rho(X_1) + c, \forall c \in \mathbb{R}\). Moreover, if we consider a second random variable \(X_2 \in L_2(\mathbb{P})\), comonotonic additivity means that \(\rho(X_1 + X_2) = \rho(X_1) + \rho(X_2)\) if \(X_1\) and \(X_2\) are comonotonic, i.e. they are perfectly positive dependent. We observe that comonotonic additivity is an important aspect of subadditivity and represents the limiting case where diversification has no effect.
that is an expected utility calculated under a modified probability measure. In other words, the distortion measure defines a change of probability and represents the more or less pessimistic view of the agent on any admissible risk level. DRMs have been applied to a variety of insurance and asset allocation problems (e.g., van der Hoek and Sherris, 2001; Hamada et al., 2006; Balbás et al., 2009).

In this paper we will focus on the coherent subset of DRMs (CDRMs, for short). To generate such a coherent measure, we request the distortion function to be well-behaved, i.e. to satisfy the properties of continuity, concavity and differentiability (see Wang et al., 1997). But there is a second level of coherence that we want the DRM to meet. Given that the distortion function is a parametric function of the risk aversion parameter, we impose also a parameter coherence. This second level of coherence is not implied by the properties a risk measure has to satisfy in order to be a CDRM. We name the subset of CDRMs which satisfies the coherence condition also with respect to the risk aversion parameter as the class of fully coherent distortion risk measures (FCDRMs).

DRMs have recently gained increasing attention due to their close relationship with the class of spectral risk measures (SRMs), introduced by Acerbi (2002, 2004). SRMs relate the risk measure to the agent’s risk aversion through a weighted average of the loss distribution quantiles, where weights depend on a risk aversion parameter. Regarding the subjective specification of the weighting function, Dowd et al. (2008) suggest a purely phenomenological approach, proposing a weighting function of the same type of a given utility function. However, they note that SRMs constructed in this way do not necessarily behave consistently with the risk aversion preferences exhibited by the chosen utility function. Consequently, the class of SRMs coincides exactly with that of CDRMs, but not with that of FCDRMs, on which we focus our attention in this paper.

In conclusion, this paper builds on the existing risk management literature as follows. We propose a novel approach to derive the OHR with futures contracts, i.e. we minimize a risk measure that depends on the agent’s preferences towards risk. The derived risk-minimizing hedge ratio is consistent with the corresponding dual problem written in term of subjective expected utility maximization. To this purpose, we introduce a subset of CDRMs—that we name as fully coherent distortion risk measures—which incorporates agent’s preferences towards risk. Our approach permits to generalize the way to construct the risk

\[3\] The representability of an (extended) real-valued function on a space of risks by means of a Choquet integral has been studied by several authors, e.g. Chateauneuf (1996), Wu and Zhou (2006) and Young (1998).

\[4\] Several applications of SRMs have been suggested in the financial literature (e.g., Adam et al. 2008; Cotter and Dowd, 2006).
aversion function (as an alternative to the phenomenological approach) and the second level of coherence clarifies the role of the risk aversion parameter attached to the weighting function. To our knowledge, the latter point has never been formalized in the literature by means of a class of risk measures satisfying such a twofold coherence. Using simulated data, we analyze the sensitivity of the fully coherent OHR to the probability distribution of the hedged portfolio returns and the risk aversion parameter. We stress the importance of choosing a FCDRM instead of a measure which does not satisfy this second degree of coherence.

The remainder of the paper is organized as follows. Section 2 presents the theoretical problem and discusses the choice of a FCDRM. Section 3 provides a numerical illustration of the OHR based on different risk measures using simulated data. Section 4 concludes.

2 The optimal problem

The agent faces the problem of estimating the optimal position to be held in futures contracts in order to hedge a given spot position at the current time. The optimization program can be expressed in terms of the minimization of a given measure of risk of the hedged portfolio returns. We will allow the agent to be risk averse; hence, she will attribute higher importance to larger losses than that she attributes to smaller losses. Formally, if we define as $R^S$ and $R^F$ the returns of the spot and the futures position, respectively, the return on the hedged portfolio is given by

$$R^h = R^S - hR^F,$$

where $h$ is the number of futures contracts to be held in order to hedge a unitary spot position.\(^7\) The hedged portfolio return in equation (1) encompasses both the case of a long spot position ($h > 0$) and a short spot position ($h < 0$). The objective of hedging is to choose the optimal relative futures position $h$ (i.e., the OHR), which has to be defined in terms of some functional of the random variable $R^h$. Therefore, the solution of the optimal program requires the agent to specify a “suitable” risk measure which properly embeds her subjective prefer-

---

\(^7\)We ignore the different sizes of the spot and the futures position. In this way, $h$ can be interpreted as the relative size of the futures positions relative to one spot position. We do not lose in generality, as the problem can be straightforwardly modified to incorporate the unequal size of each position.
2.1 Risk measures

Our discussion moves now to choosing the risk measure of the hedged portfolio returns to be minimized. In this regard, we consider the coherent subset of quantile-based risk measures (QRMs), which allows to subjectively re-weight the quantiles of the returns distribution in order to introduce a dependence on agent’s risk aversion (or pessimism).\(^8\) Given a probability space \((\Omega, \mathcal{F}, P)\) and a random variable \(X \in L_2(P)\), we introduce the definition of coherent distorted risk measure (CDRM) of \(X\).\(^9\)

**Definition 1** A coherent distorted risk measure \(\rho_{\phi, \gamma}(X)\) is defined as

\[
\rho_{\phi, \gamma}(X) = -\int_0^1 \phi_\gamma(s) q_s(X) \, ds,
\]

where \(q_s(X)\) denotes the \(s\)-quantile of \(X\) and \(\phi_\gamma(s) \in L_1([0, 1])\) is a weighting function, defined for all \(s \in [0, 1]\), such that \(\phi_\gamma(s) = \partial_s g_\gamma(s)\) for a distortion function \(g_\gamma : [0, 1] \rightarrow [0, 1]\), which is non-decreasing, continuous, concave and differentiable, with \(g_\gamma(0) = 0\) and \(g_\gamma(1) = 1\).

The function \(\phi_\gamma\) is called *admissible risk spectrum* or *risk aversion function* and assigns weights to different confidence levels. According to this definition, CDRMs correspond to SRMs whose weights equal the derivative of a concave distortion risk function (see Gzyl and Mayoral, 2006). As an example, the expected shortfall, defined as

\[
\text{ES}_{1-u}(X) = -\frac{1}{u} \int_0^u q_s(X) \, ds,
\]

can be interpreted as a CDRM with risk spectrum \(\phi(s) = \frac{1}{u} 1_{0 \leq s \leq u}\), where \(1\) denotes the indicator function over the considered probability interval. Therefore, this risk spectrum represents a risk neutral behavior of the agent within the range of relevant quantiles, as it assigns the same weight (equal to \(1/u\)) to all probability levels before \(u\), and no weight elsewhere. The expected shortfall is not a complete risk measure, since the distortion function is not strictly increasing on the interval \((u, 1]\).\(^10\) In general, the risk spectrum is a parametric

---

\(^8\)For a discussion of QRMs and their application in financial risk management literature, see Dowd and Blake (2006).

\(^9\)In what follows we will treat the random variable \(X\) as defining returns. Hence, \(X\) will be defined over the real line and due to our sign convention we will measure losses in the left (negative) tail of the probability distribution of \(X\).

\(^10\)According to Balbás et al. (2009), a *complete* distortion function is strictly increasing. This condition allows to use all information of the original loss distribution. ES is not a complete risk
function. In what follows we clarify the meaning of the risk aversion parameter \( \gamma \) in a way such to satisfy a second level of coherence.

**Remark 2** As proved by Wang (1996), spectral risk measures (SRMs) can be considered equivalent to a subset of distortion risk measures (DRMs), i.e. coherent distortion risk measures (CDRMs). Particularly, if the distortion function is defined such to generate a coherent measure, the weights in equation (2) satisfy:

1. \( \phi_\gamma(s) \geq 0 \) (nonnegativity);
2. \( \int_0^1 \phi_\gamma(s) ds = 1 \) (normalization);
3. \( \partial_s \phi_\gamma(s) \leq 0 \) (decreasingness).

The first and the second coherence conditions require that the weights are always non-negative and that they sum to 1, but the key condition is the third. In fact, property 3 requires that weights attached to higher losses are not smaller than weights attached to lower losses, and this condition reflects agent’s risk aversion. This observation allows to assert that a CDRM has the structure of a SRM, whose risk spectrum is \( \partial_s g_\gamma(s) = \phi_\gamma(s) \). Tsanakas and Desli (2003) prove that the properties of the distortion function which assure the coherence of the risk measure coincide with the necessary properties of risk spectrum. Hence, the class of CDRMs coincide with that of SRMs. Among the requested properties of the DRMs, we emphasize that the risk aversion of the decision maker is ensured by requiring a concave \( g_\gamma \) (Yaari, 1987), but this is not sufficient to reach the fully coherence, as we will show in what follows.

We focus now on the analytical form of the parametric function defining the risk spectrum. As suggested by the phenomenological approach, the weighting function should be specified starting from the agent’s utility function. Bertsimas et al. (2004) observe that it is important to choose a “convenient” utility function and then translate it into a corresponding risk spectrum of a similar form. However, this phenomenological approach to weights definition is not successful in connecting the corresponding risk measure to the utility function (Gzyl and Mayoral, 2006). In this paper we do not restrict our attention to the set of risk spectra put forward by the phenomenological approach. In fact, we define the risk spectrum moving from a distortion function, and this allows us to widen the set of possible weights. Moreover, as Dowd et al. (2008) point out, the parameter of the risk spectrum may be not consistent with the interpretation measure since the distortion function is constant on the interval \((u, 1)\), being \( g(s) = s/u \) for \( 0 \leq s \leq u \) and \( g(s) = 1 \) otherwise.
as risk aversion indicator (this happens, for example, in the case of the power utility function with $\gamma < 1$) and hence \textit{ad hoc} assigning a risk spectrum to a given utility function can lead to inconsistent results. We clarify this point introducing a second level of coherence.

In the following we list some complete and coherent distortion risk functions\footnote{A complete and coherent risk measure is said to be \textit{exhaustive}, and in this paper we will consider only exhaustive risk measures.} and the corresponding risk spectra, frequently used in finance and insurance (see Wang et al., 1997). Figure 1 plots the risk spectrum as a function of the probability level for different values of the risk aversion parameter. It is evident that all the considered risk measures respect the first level coherence property (see the characteristics of $\phi_\gamma$ in Remark 2). Figure 2 plots the different risk measures against the risk aversion parameter, $\gamma$, for some usual probability distributions modeling returns. We choose the Normal distribution as the standard example, and we separately introduce excess kurtosis (Student’s t-distribution) and left skewness (Gumbel distribution) to mimic the well-known return stylized characteristics.

1. Exponential distortion.

$$g_\gamma(s) = \frac{1 - e^{-\gamma s}}{1 - e^{-\gamma}} \rightarrow \phi_\gamma(s) = \frac{\gamma e^{-\gamma s}}{1 - e^{-\gamma}}$$

This function reflects a constant absolute risk aversion. As Dowd \textit{et al.} (2008) point out, the exponential risk measure converges to the mean of the loss distribution as the coefficient of absolute risk aversion goes to zero for any probability distribution. Figure 2 shows that the corresponding risk measure is always increasing with respect to the risk aversion parameter.

2. Power distortion.

$$g_\gamma(s) = \begin{cases} 
  s^\gamma & \gamma < 1 \\
  (1 - s)^\gamma & \gamma > 1
\end{cases} \rightarrow \phi_\gamma(s) = \begin{cases} 
  \gamma s^\gamma - 1 & \gamma < 1 \\
  \gamma (1 - s)^{\gamma - 1} & \gamma > 1
\end{cases}$$

Please insert Figure 1 about here

Please insert Figure 2 about here
Here, $\gamma > 0$ is the constant coefficient of relative risk aversion. The power risk measure approaches the mean of the underlying loss distribution as the risk aversion parameter $\gamma$ goes to 1. Figure 2 shows that the corresponding risk measure is always increasing with respect to $\gamma$ if $\gamma > 1$, while in the case $\gamma < 1$ its behavior is not monotone over the entire domain.

3. Dual-power distortion.

$$g_\gamma(s) = 1 - (1 - s)^{\gamma} \rightarrow \phi_\gamma(s) = \gamma(1 - s)^{\gamma - 1}, \quad \gamma \geq 1$$

The dual-power risk measure tends to the mean of the loss distribution as the risk aversion parameter goes to 1. As depicted by figure 2, the corresponding risk measure is always increasing with respect to $\gamma$.


$$g_\gamma(s) = \Phi \left( \Phi^{-1}(1 - s) - \gamma \right) \rightarrow \phi_\gamma(s) = \exp \left( \gamma \Phi^{-1}(1 - s) - \frac{\gamma^2}{2} \right)$$

where $\Phi$ denotes the Normal cumulative distribution function. This risk measure is not monotone with respect to $\gamma$ as it is evident in figure 2. 12

5. Proportional hazard distortion.

$$g_\gamma(s) = s^{1/\gamma} \rightarrow \phi_\gamma(s) = \frac{1}{\gamma} s^{1/\gamma - 1}, \quad \gamma \geq 1$$

The proportional hazard risk measure approaches the mean of the underlying loss distribution as the risk aversion parameter $\gamma$ goes to 1. Figure 2 depicts a non-monotone behavior of the risk measure with respect to $\gamma$.

The set of CDRMs can be seen as a class of measures based on a re-weighting of the initial distribution, with weights equal to the derivative of a concave distortion risk function or risk aversion function. As we pointed out in remark 2, this condition on the distortion function assures to recover a coherent risk measure, in the usual meaning (Artzner et al., 1997, 1999). In our next step we will stress a coherence property of the risk measure also with respect to the risk aversion parameter, $\gamma$. As figure 2 clarifies, some CDRMs exhibit a behavior inconsistent with the interpretation of $\gamma$ as risk aversion indicator, since the risk

---

12 This is the unique distortion function that allows to recover CAPM and Black-Scholes model, and the parameter $\gamma$ can be interpreted as the market price of risk and reflects systematic risk (see Wang, 2000).
associated to higher risk aversion parameters is smaller than that associated to lower degrees of risk aversion. Therefore, we restrict our attention to a subset of the class of CDRMs—that we call fully coherent distorted risk measures (FCDRM)—which satisfies an additional level of coherence, i.e. with respect to the risk aversion parameter.

**Definition 3** A FCDRM is a CDRM \( \rho_{\phi_{\gamma}}(X) \) where \( \phi_{\gamma}(s) \in L_1([0,1]) \), such that 
\[
\partial_s \phi_{\gamma}(s) \leq 0, \forall \gamma \in \mathbb{R}^+ \text{ and } \partial_{\gamma} \rho_{\phi_{\gamma}}(X) > 0.
\]

This definition requires that a CDRM satisfies a further condition, i.e. it is an increasing function of the risk aversion parameter, \( \gamma \). This requirement is intuitive but nonetheless central in order to interpret \( \rho_{\phi} \) as a risk measure. In fact, measures which exhibit a non-monotonic and increasing behavior with respect to the risk aversion parameter over the entire domain do not correctly assess risk. Obviously, CDRMs depend on the probability distribution of the underlying random variable, \( X \). In our framework \( X \) represents returns, hence we restrict our attention to the set of probability distributions which commonly model this variable.\(^{13}\)

**Lemma 4** FCDRM \( \subset \) CDRM.

**Proof.** To prove this strict inclusion it is sufficient to assume that the thesis is false and to give a counterexample. Let us consider the power distortion function with parameter \( \gamma < 1 \). This distortion function generates a complete coherent measure since 
\[
\partial_u \phi_{\gamma}(u) = \gamma(\gamma - 1)u^{\gamma - 2} < 0, \forall \gamma \in \mathbb{R}^+ \text{ (see figure 1)}.
\]
But this measure is not fully coherent as it is evident observing figure 2, since there exist a portion of the domain of \( \gamma \) in which \( \partial_{\gamma} \rho_{\phi_{\gamma}}(X) < 0 \). Then the thesis is confirmed. \( \blacksquare \)

Finally, we verify that any decision maker’s risk-averse behavior can been coded in a risk spectrum and hence modeled by a CDRM. Following Wächter and Mazzoni (2010), in appendix we clarify that this problem is consistent with the optimization of the expected utility of an agent whose preferences are modeled within the dual theory of choice.

### 2.2 Optimal hedge ratio

After choosing a proper risk measure, i.e. a risk measure which belongs to the FCDRMs class, we solve the OHR problem minimizing the risk of the hedged

\(^{13}\)Dowd et al. (2008) prove that it is not analytically feasible to show that \( \partial_{\gamma} \rho_{\phi_{\gamma}}(X) \) has a univocal sign for all possible probability distributions. Hence, we choose some usual distributions exhibiting the common empirical features of financial returns.
portfolio returns with respect to $h$. Hence, the optimal hedge ratio based on $\rho_\phi$ is the solution of the following program:

$$h^* = \arg\min_h \rho_\phi (R^h), \quad (4)$$

where $\rho_\phi (R^h)$ depends on the quantiles of $R^h$, and thus on the hedge ratio $h$.

3 Simulation

In this section we solve the optimization problem in equation (4) for the fully coherent and the non-fully coherent risk measures we have previously proposed—given some common return probability distributions—and we discuss how the estimated OHR is affected by a wide range of returns characteristics. We conduct our analysis by means of a simulation approach in order to disentangle the effect on the OHR of a change in each single parameter, leaving the others unchanged.

For our simulation we proceed as follows. Initially, (a) we extract 250 returns on the spot and the futures positions drawn from a given probability distribution with known parameters, e.g., volatility of spot returns $R^S$ and futures returns $R^F$ and their correlation. As we did in the previous section, we choose the Normal distribution, the Student’s t-distribution and the Gumbel distribution. Since these data are intended to simulate 250 daily returns, i.e. one year in terms of trading days, we assume that both $R^S$ and $R^F$ have zero expected value. Then, (b) based on simulated data for spot and futures returns we compute the return on the hedged portfolio, that is the random vector $R^h$ given by equation (1), for a given value of the hedge ratio, $h$. Finally, (c) we compute the risk of hedged portfolio returns using a FCDRM and we numerically estimate the OHR choosing the level of $h$ which minimizes the chosen risk measure. We denote this $h$ as $h^*$.14 We repeat steps from (a) to (c) 1,000 times, ending with 250,000 simulated spot and futures returns, and 1,000 estimated OHRs for each probability distribution and relevant parameters. Table 1 shows the results of our simulation for the three FCDRMs of the previous section.

14For computational matters, we limit $h$ to the interval $[0, 2]$, meaning no hedging ($h = 0$) and hedging with a futures position twice as large as the spot position ($h = 2$), respectively. These bounds are structural to our problem, as (a) since spot and futures positions are positively correlated $h$ cannot go negative, and (b) for common levels of correlation and volatility in practice $h > 2$ is far from being observed.
Panel A of table 1 shows average OHRs when both spot and futures returns are drawn from a Normal probability distribution. Different levels of correlation between spot and futures returns and volatility are taken into consideration. Very high correlation (equal to 0.95) and same volatility (almost perfect hedge) is shown in the first row; correlation equal to 0.80 and same volatility, correlation equal to 0.80 and spot volatility 1.25 times larger than futures volatility, and correlation equal to 0.80 and spot volatility 20% smaller than futures volatility are depicted in rows two, three and four, respectively. The second and third panel of table 1 carry out the same comparative statics analysis for spot and futures exhibiting excess kurtosis (Student’s t distribution, panel B) and negative skewness (Gumbel distribution, panel C). The considered risk measures include the exponential, the power with $\gamma > 1$ and the dual-power; variance and expected shortfall (at 95% confidence level) are added for comparative purposes.

As we expected, when returns are normally distributed (panel A) risk aversion has no effect on the estimation of the OHR. Results for all risk measures (and expected shortfall, which implicitly assumes risk neutrality) converge to the MVHR shown in column one, that is the correlation coefficient times the ratio between spot and futures volatility. Benninga et al. (1983) and Lien and Tse (2002) find the same result for the utility maximizing hedge ratio and this evidence confirms the commonalities of the two procedures. This is an intuitive result as the information provided by a Normal distribution is entirely embedded into its first two moments. When data are not drawn from a Normal distribution (as it is always the case in practice), risk aversion does affect the estimation of the OHR. The difference between the MVHR and the OHR obtained minimizing a FCDRM widens as data get far from the normality assumption. Gumbel distribution, which introduces left skewness (as it usually the case when considering daily returns), shows the largest differences. Also, basis risk—that can be approximated with the cases of different spot and futures return volatility, along with a low correlation of returns—amplifies this conclusion. However, a clear evidence of table 1 is that the estimated OHR is little sensitive to the chosen FCDRM. This is apparent also from inspecting figure 1, as the risk spectrum of exponential, power ($\gamma > 1$) and dual-power risk measures exhibits approximate similar shapes and values, and the same holds for the risk measures themselves (see figure 2).

Figure 3 focuses down the analysis and compares the behavior of the OHR
calculated minimizing a FCDRM, i.e. the exponential risk measure, with that obtained minimizing a non-FCDRM, i.e. proportional hazard risk measure, for the three selected probability distributions and the particular case of correlation equal to 0.80 and same volatility of spot and futures returns. Once again, the normality of returns implies that risk aversion has no effect on the OHR, since it is represented by a flat line. Departure from normality makes clear that the fully coherence characteristic of the risk measure leads to a different behavior of the OHR, which is always increasing with $\gamma$ for Student’s t-distribution while it shows a decreasing region of the domain in the Gumbel distribution case. The OHR determined through a non-FCDRM exhibits a different behavior, as it is non-monotone for Student’s t-distribution and strictly decreasing in the Gumbel distribution example.

Table 2 replicates the analysis carried out in table 1 for the three non-FCDRM presented in the previous section, that is the Wang’s, the proportional hazard and the power with $\gamma < 1$ risk measure. From the inspection of this table, and comparing the estimated OHRs with those depicted in table 1, it is evident that in this case the choice of the risk measure affects considerably the results. Especially in the Gumbel case this is clear, as for instance the power risk measure results in generally higher OHRs. This obviously depends on the fact that, contrarily to what happens in table 1, here the risk aversion parameter is not immediately comparable within different risk measures. But moreover and most important, the pattern of the OHR as a function of the risk aversion parameter exhibits a different behavior. In this sense, table 2 confirms the evidence reported by figure 3. Let’s focus our attention, as an example, on Gumbel distribution (table 2, panel C). The minimum-proportional hazard risk measure hedge ratio is a decreasing function of the risk aversion parameter, $\gamma$. In fact, for $\gamma = 10, 25, 50$ the OHR results equal to 1.486, 1.473 and 1.451, respectively, in the case of correlation 0.80 and same spot and futures returns volatility. It is clear from panel C of table 1 that for the same parameters the OHR is increasing in $\gamma$, regardless of the (fully coherent distortion) risk measure we choose. This means that if we move from a moderate degree of risk aversion to a more marked one, we have to decrease the OHR if we measure risk through a proportional hazard risk measure and increase the OHR if we measure risk by means of (for example) an exponential risk measure. The opposite behavior stems from the fact that the proportional hazard risk measure (and the same holds for the other
non-FCDRMs) is a decreasing function of the risk aversion parameter, for these values of $\gamma$. This counterintuitive behavior is clear inspecting panel C of figure 2. In other words, non-FCDRMs fail in measuring risk, and consequently fail in estimating the OHR.

4 Conclusion

Literature on the derivation of the optimal hedge ratio (OHR) with futures contracts has focused on either a risk-minimization or utility maximization framework. While the latter structurally incorporates risk aversion through agent’s expected utility, the former approach largely neglects subjective preferences. In this paper we have proposed a risk-minimizing hedge ratio where the subjective characterization of the risk measure in terms of agent’s risk aversion is made through a novel class of risk measures satisfying a twofold level of coherence, i.e. both in the meaning of Artzner et al. (1997, 1999) and with respect to the risk aversion parameter, expressing agent’s perception of losses. This class—that we have named as *fully coherent distortion risk measures* (FCDRMs)—is a subset of coherent distortion risk measures (CDRMs) to which we have imposed a further condition on the interpretation of the subjective weighting function. Employing FCDRMs we have estimated the OHR using simulated data and we have analyzed the sensitivity of the fully coherent OHR to the probability distribution of the hedged portfolio returns and the risk aversion parameter. As expected, as soon as data are not drawn from a Normal distribution, the estimated OHR significantly diverges from the minimum-variance hedge ratio (MVHR). Moreover, our results show that neglecting the second degree of coherence leads to an inconsistent behavior of the OHR, as non-FCDRMs fail in correctly measuring risk. The results confirm that, due to data non-normality, coherently weighting agent’s risk aversion plays a substantial role.

A Appendix

In this appendix we stress the fact that the minimization of a CDRM corresponds to the maximization of the expected utility in the sense of the *dual theory of choice* developed by Yaari (1987). The consistency of these minimization and maximization problems is not verified with respect to the usual expected utility theory. In fact, for a given preference relation $\succeq$, a representation in term of expected utility, that is $U(X) = \int_{-\infty}^{+\infty} u(x)dF_X(x)$, where $F_X$ is the cumulative dis-
tribution function of $X$ and $u$ is a given utility function, exists if and only if the preference relation obeys a certain set of axioms (von Neumann and Morgenstern, 1947). Among them, the independence axiom is critical for our problem,\footnote{Sriboonchitta et al. (2010) develop a calculation scheme for the construction of a risk spectrum $\phi_\gamma$ consistent with a given utility function, but they fail in establishing a relationship between expected utility and spectral risk since non-trivial applications generate inconsistent results.} i.e.

$$X \succeq Y \Rightarrow \lambda F_X(x) + (1 - \lambda) F_Z(x) \succeq \lambda F_Y(x) + (1 - \lambda) F_Z(x),$$

$$\forall X, Y, Z \in L_2(\mathcal{P}), x \in \mathbb{R}, \lambda \in [0, 1].$$

It is only considering a \textit{weaker} version of the independence-axiom, as proposed by Puppe (1991), that is possible to verify the consistency between these two problems. In an alternative decision theory called \textit{dual theory of choice} (see Yaari, 1987), the independence axiom is replaced by the following weaker requirement:

$$X \succeq Y \Rightarrow (\lambda F_X \ast (1 - \lambda) F_Z)(x) \succeq (\lambda F_Y \ast (1 - \lambda) F_Z)(x),$$

$$\forall X, Y, Z \in L_2(\mathcal{P}), x \in \mathbb{R}, \lambda \in [0, 1]$$

where $\ast$ indicates convex convolution. This new requirement postulates not the independence of the convex combination of the distributions, but the independence of the convex combination of the random variables themselves. Finally we have the following result.

**Proposition 5** Given a probability space $(\Omega, \mathcal{F}, \mathcal{P})$ and the random variables $X, Y \in L_2(\mathcal{P})$, any CDRM (or SRM) is consistent with the dual theory of choice:

$$\rho_{\phi_\gamma}(X) \leq \rho_{\phi_\gamma}(Y) \Leftrightarrow E_{g_\gamma}[X] \geq E_{g_\gamma}[Y], \quad \forall X, Y \in L_2(\mathcal{P})$$

for any $g_\gamma(u) = \int_0^u \phi_\gamma(s)ds$.

**Proof.** We want to establish a systematic relationship between the risk spectrum $\phi_\gamma$ and the \textit{dual theory of choice} such that

$$\rho_{\phi_\gamma}(X) \leq \rho_{\phi_\gamma}(Y) \Leftrightarrow \tilde{U}(X) \geq \tilde{U}(Y), \quad \forall X, Y \in L_2(\mathcal{P}),$$

(5)

where $\tilde{U}(.)$ is a function representing agent’s preferences in line with the \textit{dual theory of choice}. This theory, developed by Yaari (1987), represents preference relations through the comparison of the expected values of random variables.
computed with respect to a transformed distribution, corresponding to a change of measure. Then, the preference relation in (5) is represented as follows:

$$\tilde{U}(X) = \mathbb{E}_{g_{\gamma}}[X] = \int_{-\infty}^{+\infty} xd(g_{\gamma} \circ F_X)(x),$$

(6)

where $g_{\gamma} : [0,1] \rightarrow [0,1]$ is a non-decreasing distortion function such that $g_{\gamma}(0) = 0$ and $g_{\gamma}(1) = 1$. Then applying the premium principle,\(^{16}\) we induce the following DRM from the preference relation (6):

$$\rho_{g_{\gamma}}(X) = -\mathbb{E}_{g_{\gamma}}[X] = -\int_{0}^{1} q_s(X)dg_{\gamma}(s) = -\int_{0}^{1} q_s(X)\partial_s g_{\gamma}(s)ds,$$

which shows that $\rho_{g_{\gamma}}(X)$ has the structure of a spectral risk measure $\rho_{\phi_{\gamma}}(X)$, whose risk spectrum is $\partial_s g_{\gamma}(s) = \phi_{\gamma}(s)$. This measure satisfies the necessary coherence properties if the distortion function is concave (Tsanakas and Desli, 2003). The correspondence between SRMs and CDRMs allows to prove the thesis.

References


\(^{16}\)In order to determine a price or premium for a risk it is necessary to convert the random future gain or loss into financial terms and this is done by premium principles, which convert risk measures into monetary terms. Different premium principles have been proposed in actuarial science (see Goovaerts, 1984). Here we consider the premium principle proposed by Wang (1995) that corresponds to the certainty equivalent of the *dual theory of choice*.  

16


Figure 1 – Risk spectrum for different values of the risk aversion parameter. The figure shows the value of the exponential, power ($\gamma > 1$), dual-power, Wang, proportional hazard and power ($\gamma < 1$) risk spectrum as a function of the probability level. Each chart depicts the risk spectrum for three different values of the risk aversion parameter, $\gamma$. 
Panel A – Returns drawn from a Normal distribution.

Panel B – Returns drawn from a Student’s t-distribution.

Panel C – Returns drawn from a Gumbel distribution.

Figure 2 – Risk measures against the risk aversion parameter. The figure shows the value of the exponential, power ($\gamma > 1$), dual-power, Wang, proportional hazard and power ($\gamma < 1$) risk measures as a function of the risk aversion parameter, $\gamma$. 10,000 values are extracted and they simulate daily returns (zero expected value and annual volatility equal to 30%). In panel A returns are drawn from a Normal distribution; in panel B returns are drawn from a Student’s t-distribution; in panel C returns are drawn from a Gumbel distribution.
Figure 3 – Optimal hedge ratio against the risk aversion parameter. The figure depicts the optimal hedge ratio as a function of the risk aversion parameter, $\gamma$, for different choices of risk measure and return distribution. In particular, the solid line corresponds to the exponential risk measure and the dotted line denotes the proportional hazard risk measure. Returns are drawn from a Normal distribution (top left chart), a Student’s $t$-distribution (top right chart) and a Gumbel distribution (bottom chart). The table gives a numeric representation of the optimal hedge ratio obtained using the exponential and the proportional hazard risk measure, for the three probability distributions and four different degrees of risk aversion (viz., $\gamma = 3, 5, 15, 25$).
### Panel A: Returns drawn from a Normal distribution

<table>
<thead>
<tr>
<th>Variance</th>
<th>Expected shortfall</th>
<th>Exponential RM</th>
<th>Power RM</th>
<th>Dual-power RM</th>
</tr>
</thead>
<tbody>
<tr>
<td>OHR</td>
<td>OHR</td>
<td>γ OHR</td>
<td>γ OHR</td>
<td>γ OHR</td>
</tr>
<tr>
<td>corr = 0.95, $\sigma_S/\sigma_F = 1.0$</td>
<td>0.951</td>
<td>0.951</td>
<td>10 0.950</td>
<td>10 0.950</td>
</tr>
<tr>
<td>corr = 0.80, $\sigma_S/\sigma_F = 1.0$</td>
<td>0.800</td>
<td>0.800</td>
<td>10 0.798</td>
<td>10 0.798</td>
</tr>
<tr>
<td>corr = 0.80, $\sigma_S/\sigma_F = 1.25$</td>
<td>1.001</td>
<td>1.000</td>
<td>25 1.003</td>
<td>25 1.003</td>
</tr>
<tr>
<td>corr = 0.80, $\sigma_S/\sigma_F = 0.80$</td>
<td>0.639</td>
<td>0.638</td>
<td>25 0.641</td>
<td>25 0.641</td>
</tr>
</tbody>
</table>

### Panel B: Returns drawn from a Student’s t-distribution

<table>
<thead>
<tr>
<th>Variance</th>
<th>Expected shortfall</th>
<th>Exponential RM</th>
<th>Power RM</th>
<th>Dual-power RM</th>
</tr>
</thead>
<tbody>
<tr>
<td>OHR</td>
<td>OHR</td>
<td>γ OHR</td>
<td>γ OHR</td>
<td>γ OHR</td>
</tr>
<tr>
<td>corr = 0.95, $\sigma_S/\sigma_F = 1.0$</td>
<td>0.950</td>
<td>0.958</td>
<td>10 0.950</td>
<td>10 0.950</td>
</tr>
<tr>
<td>corr = 0.80, $\sigma_S/\sigma_F = 1.0$</td>
<td>0.802</td>
<td>0.829</td>
<td>10 0.804</td>
<td>10 0.804</td>
</tr>
<tr>
<td>corr = 0.80, $\sigma_S/\sigma_F = 1.25$</td>
<td>1.001</td>
<td>1.035</td>
<td>25 1.026</td>
<td>25 1.026</td>
</tr>
<tr>
<td>corr = 0.80, $\sigma_S/\sigma_F = 0.80$</td>
<td>0.638</td>
<td>0.660</td>
<td>25 0.654</td>
<td>25 0.654</td>
</tr>
</tbody>
</table>

### Panel C: Returns drawn from a Gumbel distribution

<table>
<thead>
<tr>
<th>Variance</th>
<th>Expected shortfall</th>
<th>Exponential RM</th>
<th>Power RM</th>
<th>Dual-power RM</th>
</tr>
</thead>
<tbody>
<tr>
<td>OHR</td>
<td>OHR</td>
<td>γ OHR</td>
<td>γ OHR</td>
<td>γ OHR</td>
</tr>
<tr>
<td>corr = 0.95, $\sigma_S/\sigma_F = 1.0$</td>
<td>0.946</td>
<td>1.228</td>
<td>10 1.190</td>
<td>10 1.188</td>
</tr>
<tr>
<td>corr = 0.80, $\sigma_S/\sigma_F = 1.0$</td>
<td>0.790</td>
<td>1.354</td>
<td>10 1.272</td>
<td>10 1.269</td>
</tr>
<tr>
<td>corr = 0.80, $\sigma_S/\sigma_F = 1.25$</td>
<td>0.985</td>
<td>1.689</td>
<td>10 1.591</td>
<td>10 1.587</td>
</tr>
<tr>
<td>corr = 0.80, $\sigma_S/\sigma_F = 0.80$</td>
<td>0.633</td>
<td>1.093</td>
<td>10 1.022</td>
<td>10 1.019</td>
</tr>
</tbody>
</table>

Table 1 – Average optimal hedge ratios for different probability distributions and parameters using a fully coherent risk measure. The table shows the numerically estimated average optimal hedge ratio obtained solving the problem in equation (7) when spot and futures returns are drawn from a Normal distribution (panel A), a Student’s t-distribution (panel B) and a Gumbel distribution (panel C). 1,000 optimal hedge ratios are estimated using a 250-day window. Employed risk measures are the exponential risk measure, the power ($\gamma > 1$) risk measure and the dual-power risk measure, for three different degrees of risk aversion, viz. $\gamma = 10, 25$ and $50$. Minimum-variance hedge ratio and minimum-expected shortfall (at 95% confidence level) hedge ratio are added for comparative purpose. Four cases are considered for each probability distribution, depending on the correlation between $R^S$ and $R^F$ (0.95 and 0.80) and the ratio between the volatility of $R^S$ and $R^F$ (1.00, 1.25 and 0.80).
### Panel A: Returns drawn from a Normal distribution

<table>
<thead>
<tr>
<th>Variance</th>
<th>Expected shortfall</th>
<th>Wang's RM</th>
<th>Proportional hazard RM</th>
<th>Power RM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OHR</td>
<td>$\gamma$</td>
<td>OHR</td>
<td>$\gamma$</td>
</tr>
<tr>
<td>corr = 0.95, $\sigma_S/\sigma_F = 1.0$</td>
<td>0.950</td>
<td>0.950</td>
<td>1.0</td>
<td>0.951</td>
</tr>
<tr>
<td>corr = 0.80, $\sigma_S/\sigma_F = 1.0$</td>
<td>0.801</td>
<td>0.801</td>
<td>1.0</td>
<td>0.802</td>
</tr>
<tr>
<td>corr = 0.80, $\sigma_S/\sigma_F = 1.25$</td>
<td>1.001</td>
<td>1.000</td>
<td>2.5</td>
<td>0.897</td>
</tr>
<tr>
<td>corr = 0.80, $\sigma_S/\sigma_F = 0.80$</td>
<td>0.640</td>
<td>0.639</td>
<td>2.5</td>
<td>0.639</td>
</tr>
</tbody>
</table>

### Panel B: Returns drawn from a Student's t-distribution

<table>
<thead>
<tr>
<th>Variance</th>
<th>Expected shortfall</th>
<th>Wang's RM</th>
<th>Proportional hazard RM</th>
<th>Power RM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OHR</td>
<td>$\gamma$</td>
<td>OHR</td>
<td>$\gamma$</td>
</tr>
<tr>
<td>corr = 0.95, $\sigma_S/\sigma_F = 1.0$</td>
<td>0.951</td>
<td>0.958</td>
<td>2.5</td>
<td>0.958</td>
</tr>
<tr>
<td>corr = 0.80, $\sigma_S/\sigma_F = 1.0$</td>
<td>0.797</td>
<td>0.823</td>
<td>1.0</td>
<td>0.800</td>
</tr>
<tr>
<td>corr = 0.80, $\sigma_S/\sigma_F = 1.25$</td>
<td>0.999</td>
<td>1.039</td>
<td>2.5</td>
<td>1.047</td>
</tr>
<tr>
<td>corr = 0.80, $\sigma_S/\sigma_F = 0.80$</td>
<td>0.640</td>
<td>0.660</td>
<td>2.5</td>
<td>0.666</td>
</tr>
</tbody>
</table>

### Panel C: Returns drawn from a Gumbel distribution

<table>
<thead>
<tr>
<th>Variance</th>
<th>Expected shortfall</th>
<th>Wang's RM</th>
<th>Proportional hazard RM</th>
<th>Power RM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OHR</td>
<td>$\gamma$</td>
<td>OHR</td>
<td>$\gamma$</td>
</tr>
<tr>
<td>corr = 0.95, $\sigma_S/\sigma_F = 1.0$</td>
<td>0.945</td>
<td>1.230</td>
<td>1.0</td>
<td>1.273</td>
</tr>
<tr>
<td>corr = 0.80, $\sigma_S/\sigma_F = 1.0$</td>
<td>0.794</td>
<td>1.350</td>
<td>1.0</td>
<td>1.435</td>
</tr>
<tr>
<td>corr = 0.80, $\sigma_S/\sigma_F = 1.25$</td>
<td>0.987</td>
<td>1.701</td>
<td>2.5</td>
<td>1.706</td>
</tr>
<tr>
<td>corr = 0.80, $\sigma_S/\sigma_F = 0.80$</td>
<td>0.631</td>
<td>1.076</td>
<td>2.5</td>
<td>1.092</td>
</tr>
</tbody>
</table>

**Table 2** – Average optimal hedge ratios for different probability distributions and parameters using a non-fully coherent risk measure. The table shows the numerically estimated average optimal hedge ratio obtained solving the problem in equation (7) when spot and futures returns are drawn from a Normal distribution (panel A), a Student’s t-distribution (panel B) and a Gumbel distribution (panel C). 1,000 optimal hedge ratios are estimated using a 250-day window. Employed risk measures are the Wang’s risk measure, the proportional hazard risk measure and the power ($\gamma < 1$) risk measure, for three different degrees of risk aversion, viz. $\gamma = 1.0, 2.5$ and 5.0 (Wang), $\gamma = 10, 25$ and 50 (proportional hazard) and $\gamma= 0.10, 0.25$ and 0.50 (power). Minimum-variance hedge ratio and minimum-expected shortfall (at 95% confidence level) hedge ratio are added for comparative purpose. Four cases are considered for each probability distribution, depending on the correlation between $R^S$ and $R^F$ (0.95 and 0.80) and the ratio between the volatility of $R^S$ and $R^F$ (1.00, 1.25 and 0.80).