AESTIMATIO, THE IEB INTERNATIONAL JOURNAL OF FINANCE, 2014. **8**: 184-201

Measuring market risk under the Basel accords: *VaR*, stressed *VaR*, and expected shortfall

Chen, James Ming

RECEIVED: 4 DECEMBER 2013
ACCEPTED: 19 MARCH 2014

Abstract

Each of the most recent accords of the Basel Committee on Banking Regulation, known as Basel II, 2.5, and III, has embraced a different primary measure of market risk in global banking regulation: traditional value-at-risk (*VaR*), stressed *VaR*, and expected shortfall. After introducing the mathematics of VaR and expected shortfall, this article will evaluate how well the reforms embraced by Basel 2.5 and III — stressed VaR and expected shortfall — have addressed longstanding regulatory concerns with traditional *VaR*. Expected shortfall, but not *VaR*, provides a coherent measure of risk. On the other hand, *VaR*, but not expected shortfall (or, for that matter, nearly every other general spectral measure of risk), satisfies the mathematical requirement of "elicitability." Mathematical limitations on measures of risk therefore force regulators and bankers to choose between coherence and elicitability, between theoretically sound consolidation of diverse risks and reliable backtesting of risk forecasts against historical observations.

Keywords:

Value-at-risk, stressed *VaR*, expected shortfall, coherence, elicitability, Basel.

JEL classification:

G15, G38, G32

Chen, J. M. ✉ *Justin Smith Morrill Chair in Law, Michigan State University College of Law, 648 North Shaw Lane, East Lansing, Michigan 48824-1300, USA.* ✆ *(502) 432-6891 (502) 432-6801(fax). Email: chenjame@law.msu.edu*

AESTIMATIO, THE IEB INTERNATIONAL JOURNAL OF FINANCE, 2014. **8**: 184-201

Medición de riesgo de mercado según los acuerdos de Basilea: ***VaR, VaR* en situaciones de estrés y pérdida esperada**

Chen, James Ming

Resumen

Cada uno de los recientes acuerdos del Comité de Basilea sobre Regulación Bancaria, conocidos como Basilea II, 2.5 y III ha adoptado una medida primaria de riesgo en el ámbito de la regulación bancaria global: el tradicional valor en riesgo (*VaR*), el *VaR* en situaciones de estrés y la pérdida esperada. Tras introducir los aspectos cuantitativos del *VaR* y la pérdida esperada, en este artículo se evalúa el grado en el que las reformas adoptadas por Basilea 2.5 y III— *VaR* en situaciones de estrés y pérdida esperada — han abordado las tradicionales preocupaciones regulatorias relativas al *VaR*. La pérdida esperada, que no el *VaR*, proporciona una medida coherente del riesgo. Por otra parte, el *VaR*, que no la pérdida esperada (o, para el caso, casi cualquier otra medida de riesgo espectral general), satisface los requisitos matemáticos de "elicitabilidad". Las limitaciones matemáticas en las medidas de riesgo fuerzan, por tanto, a reguladores y banqueros a elegir entre coherencia y elicitabilidad, entre la consolidación teóricamente adecuada de diversos riesgos y las pruebas retrospectivas teóricamente fiables de previsiones de riesgo en base observaciones históricas.

Palabras clave:

Valor en riesgo, *VaR* en situaciones de estrés, pérdida esperada, coherencia, elicitabilidad, Basilea.

1. Introduction

This article analyzes the measurement of market risk in the trading book of financial institutions subject to the most recent accords of the Basel Committee on Banking Regulation, known as Basel II, 2.5, and III. Each of these three accords on global banking regulation has embraced a different primary measure of market risk: traditional value-at-risk (*VaR*), stressed *VaR*, and expected shortfall. After introducing the mathematics of *VaR* and expected shortfall, this note will evaluate how well the reforms embraced by Basel 2.5 and III — stressed *VaR* and expected shortfall — have addressed longstanding regulatory concerns with traditional *VaR*.

Section 2 describes the calculation of *VaR* in its conventional form. For illustrative purposes, section 2 will describe parametric *VaR* on a Gaussian distribution. Section 3 summarizes known weaknesses in *VaR*, from inherent model and estimation risk to *VaR's* failure to perform under extreme economic stress and *VaR*'s failure to satisfy the theoretical constraints on "coherent" measurements of risk. Section 4.1 describes how to calculate expected shortfall as an extension of conditional *VaR*. It further describes how expected shortfall, but not *VaR*, provides a coherent measure of risk. Section 4.2 then reverses field. It explains how *VaR*, but not expected shortfall (or, for that matter, nearly every other general spectral measure of risk), satisfies the mathematical requirement of "elicitability." Mathematical limitations on measures of risk therefore force regulators and bankers to choose between coherence and elicitability, between theoretically sound consolidation of diverse risks and reliable backtesting of risk forecasts against historical observations.

2. Conventional value-at-risk analysis

Like modern portfolio theory and the entire edifice of quantitative finance derived from those beginnings (Allen *et al.*, 2004; Benninga and Wiener, 1998; Jorion, 2006; and Mina and Xiao, 2001), conventional value-at-risk analysis assumes that risk is randomly distributed, not correlated (Whitehead, 2011). Despite its flaws and limitations (Macchiarola, 2009), *VaR* analysis arguably represents the most important tool for evaluating market risk as one of several threats to the global financial system. Basel II identifies a version of *VaR* analysis as that accord's preferred tool for assessing banks' exposure to market risk (Basel Committee, 2004). Authorities around the world have endorsed *VaR*, either as a regulatory standard or as a best practice (Federal Deposit Insurance Corporation, 1999; European Commission, 2014). Even absent regulatory compulsion, private firms routinely use *VaR* as an internal risk management tool, often directing traders to reduce exposure below the level prescribed by those firms' own *VaR* limits (Whitehead, 2011).

Let us begin by conducting an exercise in parametric *VaR*, the simplest version of *VaR* (Allen *et al*., 2004, and Jorion, 2006). Suppose that an investor stakes $1 million on an index fund tracking the Standard & Poor's 500. She asks her financial advisor, *"If capital markets go down to an extent witnessed only once in a hundred trading days, what can I lose by tomorrow's market close?"*

In its simplest form, parametric *VaR* assumes normally distributed returns (Österreichische Nationalbank, 1999). In other words, parametric *VaR* relies on the mathematics of the Gaussian distribution. Moreover, *VaR* often relies on strictly historical data (Dowd, 1998).

To answer our investor's question, an advisor using conventional parametric *VaR* analysis may assume a mean daily return of 0, with a standard deviation over that interval of 100 basis points (equal to 1 percent). On those assumptions, that advisor will report a one-day value of $VaR_{1\%}$ as $23,260 for a $1 million portfolio. $VaR_{1\%}$ = $23,260 is a fancy, technocratic way of telling this investor that she faces a 1 percent chance of losing $23,260 or more on her S&P 500 index fund on any given trading day. Equivalently, the advisor could tell the investor client that her portfolio has a 99 percent chance after a single trading day of being worth at least $976,740 ($1,000,000 – $23,260).

In formal terms, *VaR* for a certain risk or confidence level is the quantile that solves the following equation (Daníelsson and Zigrand, 2006):

$$\varepsilon = \int_{-\infty}^{-VaR} f(x)dx \quad , \tag{1}$$

where ε represents the confidence level. In the case of the investor with a $1 million portfolio tracking the S&P 500 index, $\varepsilon = 1 - .01$, or .99. $f(x)$ refers to the probability density function — in this case, of the distribution of returns on the S&P 500-indexed portfolio. *VaR* may also be defined as the greatest lower bound (infimum) on the cumulative distribution function F of any financial position Y, expressed as a real-valued, random variable (Ziegel, 2013):

$$VaR_{\alpha}(Y) = -\inf\{x \in \mathbb{R} \,|\, F_Y(x) \geq \alpha\} \quad . \tag{2}$$

Parametric *VaR* analysis requires the computation of statistical quantiles. The quantile function of a distribution is the inverse of its cumulative distribution function. As such, the quantile function is designated by the inverse of the capital phi symbol that designates the cumulative distribution function: $\Phi^{-1}(p)$. The quantile function of the standard normal distribution, also known as the probit function, is expressed as a transformation of the inverse error function:

$$z_{\alpha} = \Phi^{-1}(\alpha) = \sqrt{2}.erf^{-1}(2\alpha - 1) \ , \tag{3}$$

where *erf* refers to the error function of the normal, Gaussian distribution and erf^{-1} refers to the inverse error function.

Conventional notation in *VaR* analysis designates the quantile function as z_α. The four alternate ways for referring to the same mathematical concept — quantile function, inverse cumulative distribution function, probit, and z_α — may, somewhat surprisingly, give an affirmative, intuitive boost to the understanding of the quantitative mechanics at work. Formally, the quantile z_p represents the value at which a standard normal random variable X has exact probability p to fall inside the interval $(-\infty, z_p]$. In effect, we are asking what standard score, or z, corresponds to the value of the cumulative distribution function representing a certain percentage of the total under the curve that defines the probability density function of the returns on an investment.

We have now assembled the tools needed to compute *VaR*. Recall that we have assumed our investor has staked \$1 million in an S&P 500 index fund, where mean daily return (μ) is 0 and the standard deviation of that mean return (σ) is 100 bps (.01). The variable VaR_α expresses the value at risk given a particular probability of a loss as the product of $-z_\alpha$, standard deviation σ, and the total value of the portfolio (v) (Allen *et al.*, 2004):

$$VaR_\alpha = -z_\alpha \cdot \sigma \cdot v \quad , \tag{4}$$

The negative sign before $-z_\alpha$ allows us to state value at risk as a positive sum at risk of loss. For σ = 100 *bps* and v = \$1,000,000:

$$VaR_{.01} = -z_{.01} \cdot \sigma \cdot v \quad . \tag{5}$$

$$VaR_{.01} = -z_{.01} \cdot 100\ bps \cdot \$1{,}000{,}000 \quad . \tag{6}$$

So far we have omitted any consideration of time. As long as returns are independent and identically distributed (a crucial assumption of any distribution obeying the central limit theorem), *"variances are additive over time, which implies that volatility grows with the square root of time"* (Jorion, 2006, p. 108). To account for variance over time, we typically multiply *VaR* by the square root of time (Jorion, 2006):

$$VaR_\alpha = -z_\alpha \cdot \sigma \cdot v \cdot \sqrt{t} \quad . \tag{7}$$

Regulatory *VaR* typically assumes a ten-day holding period (equivalent to two weeks, each consisting of five trading days), which requires the multiplication of one-day *VaR* by the square root of 10, approximately 3.162. Basel II, Basel III, and the regulations of the Federal Reserve Bank of the United States all prescribe the computation of VaR on a holding period of ten business days.

All that stands between us and a complete calculation of $VaR_{.01}$ is the value of $z_{.01}$. That value in turn requires the application of the quantile function:

$$z_{.01} = \Phi^{-1}(.01) = \sqrt{2} \cdot erf^{-1}(2 \cdot .01 - 1) \approx -2.326 \quad (8)$$

Inserting this value of $z_{.01}$ into the formula for $VaR_{.01}$ yields the conclusion that $VaR_{.01}$ for this asset, over a trading interval of a single day, is approximately $23,260. The following table expresses cumulative probabilities for the foregoing exercise in parametric *VaR* at commonly used intervals (Allen *et al*., 2004):

Table 1. **Parametric *VaR* calculations at common confidence intervals**

α	.1%	.5%	1.0%	2.5%	5.0%	10%
z_α	–3.090	–2.576	–2.326	–1.960	–1.645	–1.282
VaR_α	$30,900	$25,760	$23,260	$19,600	$16,450	$12,820
$VaR_{\alpha, 10}$	$97,710	$81,460	$73,550	$61,980	$52,020	$40,540

3. Known weaknesses of conventional *VaR*

Value-at-risk analysis, especially in the simplest parametric implementation illustrated in section 2, is riddled with vulnerabilities. Three of conventional *VaR*'s weaknesses figure prominently in the Basel accords' ongoing reevaluation of measurements of risk in global financial institutions' trading books. First, epistemological limitations on our ability to describe and forecast risk undermine *VaR* as an econometric model. Second, even where the model is generally accurate, we may not supply accurate parameters. These shortcomings support a broad consensus within the financial industry that conventional *VaR* has fared poorly during periods of market stress. Finally, *VaR* does not satisfy the theoretical rigors demanded of "coherent" measures of risk.

3.1. Model and estimation risk

Like any other econometric technique, *VaR* is subject to both model risk and estimation risk (Sheppard, 2013). Reliance on a normal, Gaussian distribution systematically understates market risk borne by any portfolio. The very *"simplicity of VaR measures"*, part of this methodology's appeal to quantitative analysts and to regulators, *"is in large part obtained with assumptions not supported by empirical evidence"* (Allen *et al*., 2004, p. 8]. Of these assumptions, the *"most important (and most problematic) ... is that returns are normally distributed"* (*ibid*.). Whatever else they do, stock market returns do not follow the normal distribution (Páfka and Kondor, 2001). Departures from expected value, especially in the tails of a distribution, put a premium on statistical robustness — the resistance of a statistical model to outliers or other deviations from the model's underlying assumptions (Huber, 1981; Portnoy and He, 2000).

One quick, easily implemented way to heighten the robustness of *VaR* analysis in anticipation of unobservable and therefore unpredictable tail risks is to recalibrate parametric *VaR* according to a more leptokurtic distribution. Chen (2012) demonstrates how parametric *VaR* on a Gaussian model can be recalibrated according to a more leptokurtic distribution, such as the logistic distribution. The use of the logistic distribution, or any other elliptical distribution, merely demonstrates the fragility of the Gaussian distribution to leptokurtosis. Deft application of nonelliptical distributions such as the three-parameter lognormal (Atchison and Brown, 1957; Forsey, 2001; Limpert *et al*., 2001) or the log-logistic/Fisk distribution (Fisk, 1961) would simultaneously account for skewness and excess kurtosis in the observed distribution of returns in ways that the Gaussian distribution cannot.

Another method, embraced in the Basel accords and throughout the law of financial institutions, is to apply arbitrary multipliers such as three. For instance, one model for improving the robustness of *VaR* prescribes three states of preparedness based on multiples of *VaR*: 1 through 3 for normal conditions, 3 through 10 for stress testing, and 10+ for all other means of hedging or insuring against contingencies beyond realistic business planning (Brown, 2007).

But there may be no quantifiably reliable way of adopting a sufficiently conservative model of risk. No historic model of economic risk can predict extreme tail events. The record of monthly fluctuations in American stock market prices from 1871 through 2010 has reported 10σ events in both directions (Nordhaus, 2011), even though parametric *VaR* based on a Gaussian distribution would surmise (quite erroneously) that a 5σ event happens once every 4,000 years. A 50-year survey of oil prices, from 1960 through 2010, has revealed a 37σ event in 1973 (*ibid.*). 37σ! It would not have been unreasonable for an oil trader to believe that "the economic world as we knew it was coming to an end" (*ibid.*, p. 243). As we probe ever deeper risks and seek ever higher confidence levels, we discover to our dismay how *"we lack good analytic techniques for quantifying total risk when the distribution has a fat tail"* (Farber, 2011, p. 927).

Moreover, none of these techniques addresses two further shortcomings of *VaR*. As section 1 demonstrated, *VaR* relies on the greatest lower bound on an arbitrarily defined risk frontier over an arbitrarily fixed period of time. Basel II prescribes *VaR* at a 99 percent confidence interval over a holding period of ten trading days. Because *VaR* relies on a simple quantile analysis, it necessarily disregards the magnitude and distribution of risks in the tail beyond the designated quantile boundary (Hull, 2012). And even if we have properly modeled risk, whether by engaging in thorough nonparametric *VaR* or by specifying a more accurate parametric model of value at risk, we cannot eliminate the problem of straightforward mistakes in estimation (Sheppard, 2013).

3.2. Stressed *VaR*: Fragility during times of economic stress

As the economic crisis of 2008-09 painfully demonstrated, *VaR* fares poorly during periods of market stress. As befits an econometric model whose roots assume a symmetrical distribution of returns, *VaR* also proves fragile during periods of prosperity. Indeed, its very elegance gives rise to a false sense of security. *VaR* is at once easy to misunderstand and dangerous when misunderstood (Nocera, 2009). Tail risk poses problems during crisis and during prosperity. Because tail risk, scientifically speaking, defies measurement, the mere presence of a concrete, quantified figure such as *VaR* invites risk-taking wholly unwarranted by the real but unknown (and unknowable) state of economic affairs. Psychological anchors provided by *VaR* backfire when risks within the unobserved tail eventually materialize. In 2009 testimony urging Congress to ban *VaR*, Nassim Nicholas Taleb criticized both the scientific uncertainty of *VaR* and its psychological effects on traders (Taleb, 2009). Within the financial industry itself, David Einhorn of Greenlight Capital has echoed these criticisms, alleging that *VaR* creates perverse incentives to take *"excessive but remote risks"* and is *"potentially catastrophic when its use creates a false sense of security among senior executives and watchdogs"* (Einhorn, 2008, p. 10).

Basel 2.5 addressed some of these concerns. Basel 2.5 added "*stressed VaR*" as one of four elements in its market risk framework (Bank for International Settlements, 2011). Basel 2.5 addressed other dimensions of risk by requiring (1) an incremental risk charge for default and credit migration risk, (2) new charges for securitization and resecuritization positions within the trading book, and (3) a comprehensive risk measure for default and migration risk arising from correlation trading.

For its part, stressed *VaR* subjects conventional *VaR*, tested at the 99 percent confidence level ($1-\alpha$, where $\alpha = .01$) and with a ten-day holding period, to a one-year historic dataset that encompasses *"a continuous 12-month period of significant financial stress"* (*ibid.*, p. 2). Specifically, Basel 2.5 intended the *"stressed VaR charge ... to deliver a capital charge based on a measure of VaR that would be applicable to the bank's current portfolio in a period of stress relevant to that portfolio."* The Basel Committee's interpretative statement recognized that *"[i]n principle, the easiest way to do this is to run the current VaR model based on historical data from a period of financial stress."*

Basel 2.5 did recognize *"two particular cases where this might be inappropriate."* First, when *"a period of financial stress ... corresponds to directional moves which would lead to the* *bank making money,"* Basel 2.5 recommended the application of *"risk factor movements in both the direction which is indicated by the historical data, and the opposite direction (anti-thetic) to ensure that the period of high volatility becomes more relevant to the bank's portfolio."* Second, Basel 2.5 recognized that periods of stress may cause *"some price factors"* such as credit spreads *"to have higher absolute values"* and may distort the correspondence between large, volatile movements in those factors and *"significant increases in relative volatility (i.e. because the absolute level of the parameter is also higher)."* Under those circumstances, the Basel Committee recommended that banks *"consider modifying [their] VaR model[s] to account for large absolute factor moves that can occur in times of stress,"* as distinct from *"benign periods ... when the absolute values of credit spreads are smaller."*

Experience since Basel 2.5 has validated the reservations expressed by the Basel Committee itself. With respect to certain portions of a diverse trading portfolio, even the experience of the 2008-09 financial crisis may not properly *"represent[] a period of significant financial stress relevant to the bank's portfolio."* Certain asset classes — most prominently, perhaps, eurozone sovereign debt — experienced severe downward pressure on a timeframe distinct from that of other portions of global (or European) banking's collective balance sheet. Regulatory backtesting requirements have trouble keeping pace with the creativity (or perversion) of the financial services industry, to say nothing of exogenous forces giving rise to genuine surprises in global finance. The fundamental rule of financial regulation holds true: neither past performance nor past crisis provides any guarantee of future performance.

3.3. The incoherence of *VaR*

The final criticism of *VaR* strikes deeply at this measure's ability to describe and forecast risk. Wholly apart from methodological limitations that direct *VaR* to ignore the size and distribution of risks in the tail beyond a stipulated confidence interval, *VaR* behaves very erratically when banks or regulators try to aggregate risks associated with different components of a portfolio.

The aggregation of distinct assets or subportfolios can generate either of two basic errors. First, the aggregation of risks forecast by *VaR* can generate a type 1 error, or a *"false positive,"* by *overstating* the risk of the entire portfolio relative to those of its components. A false positive leads a bank to take less risk (and to reap less reward) than it might otherwise under the guidance of a properly calibrated measure of risk.

Second, *VaR* may *understate* the bank-wide level of risk as it aggregates risks from all branches and aspects of a bank's operations. The classic example of this type 2 error, or *"false negative,"* is the purported inability of *VaR* to express the overall risk of bank robbery. The risk that any single branch will be robbed falls well beyond the confidence level required by ordinary risk regulation. But aggregated across all branches of a large, diverse bank, the risk of at least one robbery within a manageable time frame becomes sufficiently large to fall within the confidence level and, therefore, to merit the bank's attention as well as that of its regulators.

Mathematically cogent aggregation of risk plays a crucial role in a leading theory on risk. The theory of *coherence* requires that a measure of risk satisfy four mathematical criteria: translation (or drift) invariance, (linear) homogeneity, monotonicity, and subadditivity (Sheppard, 2013; Artzner *et al*., 1999). ρ, a measure of risk, as applied to portfolios P, P_1, and P_2, is *coherent* if and only if it satisfies all of these conditions:

1. Translation (drift) invariance: Adding a constant return c to total portfolio return will reduce risk (and presumably the required regulatory capital reserve) by the same amount of c.

$$\rho(P+c) = \rho(P)-c \; . \tag{9}$$

2. (Linear) homogeneity: Increasing the size of any portfolio by a positive factor λ requires a corresponding, linear increase in regulatory capital by factor λ:

$$\rho(\lambda P) = \lambda\rho(P); \;\; \lambda > 0 \; . \tag{10}$$

3. Monotonicity: If portfolio P_1 is first-order stochastically dominant (FOSD) to portfolio P_2, in the sense that P_1 offers higher returns than P_2 in every conceivable economic state, then the risk associated with P_1 cannot be higher than P_2, and the regulatory capital required of P_1 must therefore be less than or equal to the regulatory capital required of P_2.

$$F_{P1}(x) \geq F_{P2}(x); \;\; \rho(P_1) \leq \rho(P_2) \; . \tag{11}$$

If P_1 FOSD P_2, in the formal sense that $F_{P1}(x) \geq F_{P2}(x)$ (the cumulative distribution function for P_1 is greater than or equal to the cumulative distribution function for P_2 for all values of x), then $\rho(P_1) \leq \rho(P_2)$.

4. Subadditivity: The risk associated with two combined portfolios (and the capital reserve required for those combined portfolios) cannot exceed the total risk and required capital reserve associated with each constituent portfolio, considered alone.

$$\rho(P_1+P_2) \leq \rho(P_1)+\rho(P_2) \; . \tag{12}$$

Measuring market risk under the Basel accords: *VaR*, stressed *VaR*, and expected shortfall. *Chen, James Ming*
AESTIMATIO, THE IEB INTERNATIONAL JOURNAL OF FINANCE, 2014. 8: 184-201

The first three conditions — translation invariance, linear homogeneity, and monotonicity — have proved neither difficult to satisfy nor controversial among experts in quantitative finance. By contrast, *VaR* fails to satisfy subadditivity. The following stylized example illustrates how *VaR* levels, calculated for separate portfolios, provide theoretically incoherent and practically misleading predictions of *VaR* for a combined portfolio reflecting all of its constituent components (Hull, 2012, p. 189).

Let us suppose that a bank oversees two independent projects, each with a horrifying loss profile. During the coming year, each project faces a .02 probability of a $10 million loss and a .98 probability of a $1 million loss. The one-year, 97.5% *VaR* for each project is $1 million (since the $10 million loss resides in the 2% tail beyond the 97.5% quantile of interest in this illustration). Combining the projects into the same portfolio generates the following risk profile:

- A .0004 probability of a $20 million loss (.02 x .02)
- A .0392 probability of an $11 million loss (2 x .02 x .98)
- A .9604 probability of a $2 million loss (.98 x .98)

The one-year, 97.5% *VaR* for the combined portfolio is $11 million. Again, a 2.5% *VaR* calculation predicts the loss encountered at the 2.5% quantile of the loss distribution. The predicted loss at this quantile is $11 million. But that figure grossly exceeds the combined one-year, 97.5% *VaR* values for the two projects, considered on their own: $11 million > $2 million. This violates the subadditivity condition.

VaR's failure to provide a coherent measure of risk has more than trivial consequences. The subadditivity condition suggests that a larger, combined portfolio cannot be riskier in the aggregate than the two portfolios standing apart. Except in those rare instances where a combined portfolio is so large that the unwinding of its considerable positions affirmatively affects the bid-ask spread on those assets and thereby gives rise to a problem of endogenous liquidity (Bangia *et al.*, 1998; Bervas, 2006), the aggregation of two portfolios should provide some diversification benefit and thereby lower overall risk. Since *"[s]ubadditivity reflects the idea that risk can be reduced by diversification,"* the *"use of non-subadditive risk measures"* such as *VaR* *"in a Markowitz-type portfolio optimization problem may"* invite banks to build *"portfolios that are very concentrated and that would be deemed quite risky by normal economic standards"* (McNeil *et al.*, 2005, p. 240). Even where the constituent subportfolios have a correlation of 1 and therefore provide no diversification whatsoever, a coherent risk measure would report the risk of the combined portfolio as the combined risk of each separate subportfolio:

$$\rho(P_1+P_2) = \rho(P_1)+\rho(P_2)\ . \tag{13}$$

Moreover, subadditivity improves bank regulation in other ways wholly apart from diversification benefits. *"If a regulator uses a non-subadditive risk measure in determining ... regulatory capital"* requirements, an affected financial institution *"has an incentive to legally break up into various subsidiaries in order to reduce its regulatory capital requirements"* (*ibid.*). By contrast, regulation with subadditive risk measures enables banks to decentralize their systems for risk management:

> Consider as an example two trading desks with positions leading to losses L_1 and L_2. Imagine that a risk manager wants to ensure that $R(L)$, the risk of the overall loss $L = L_1+L_2$, does not exceed some number M. If he uses a subadditive risk measure R, he may simply choose bounds M_1 and M_2 such that $M_1+ M_2 \leq M$ and impose on each of the desks the constraint that $R(L_i) \leq M_i$; subadditivity then ensures automatically that $R(L) \leq M_1+M_2 \leq M$ (*ibid.*).

4. Basel III and expected shortfall

4.1. Expected shortfall is a subadditive and coherent risk measure

Basel III appears poised to replace *VaR* with an alternative, mathematically related measure of risk, expected shortfall. In its 2011 review of academic literature concerning risk measurement, the Committee on Banking Supervision acknowledged the incoherence of *VaR* as a risk measurement (Basel Committee, 2011). Among other possibilities, the Basel Committee identified expected shortfall as a theoretically coherent alternative to *VaR*. In its May 3, 2012, consultative document on the third Basel accord, the Committee explicitly raised the prospect of phasing out *VaR* and replacing it with expected shortfall (Basel Committee, 2012).

The expected shortfall for any loss function L with confidence level $1-\alpha$ is defined formally as a transformation of VaR_α for L (Ziegel, 2013):

$$\mathrm{ES}_\alpha(L)=\frac{1}{\alpha}\int_0^\alpha VaR_\tau(L)d\tau \quad . \tag{14}$$

If L is a continuous loss distribution, then expected shortfall may be even more intuitively expressed as conditional *VaR* or the tail conditional expectation: the expected loss conditional on the loss lying beyond the limit defined by α (Basel Committee, 2011):

$$\mathrm{ES}_\alpha(L)= \mathrm{E}(L|L\geq VaR_\alpha) \quad . \tag{15}$$

As the sum of all losses exceeding the *VaR* quantile α, expected shortfall accounts for losses beyond the confidence interval. Expected shortfall is not only subadditive and coherent; it is a special case of the entire class of coherent risk measures known as spectral risk measures (Acerbi and Tasche, 2002).

A direct comparison of *VaR* and expected shortfall for the same loss distribution will illustrate how expected shortfall, unlike *VaR*, is subadditive and therefore coherent. Recall our example of two loss-laden projects managed by a bank (Hull, 2012, p. 190). Within one year, each project faces a .02 probability of a \$10 million loss and a .98 probability of a \$1 million loss. To calculate expected shortfall for each project at the 97.5 percent confidence level, observe that distribution of losses within the 2.5% tail. The worst 2 percent of this tail corresponds to a \$10 million loss; the next 0.5%, to a \$1 million loss. On the condition that we are in the 2.5% tail of the loss distribution, there is a .80 probability of a \$10 million loss and a .20 probability of a \$1 million loss. Expected shortfall at the 97.5% confidence interval for each project is therefore .8 x \$10 million + .2 x \$1 million = \$8.2 million.

Recall the loss distribution when the two projects are combined:

- A .0004 probability of a \$20 million loss (.02 x .02)
- A .0392 probability of an \$11 million loss (2 x .02 x .98)
- A .9604 probability of a \$2 million loss (.98 x .98)

Within the 2.5% tail of this loss distribution, the first .04% corresponds to a loss of \$20 million. The next 2.46 percent, to say nothing of losses beyond this boundary, corresponds to a loss of \$11 million. (We can safely ignore all other losses, including the \$2 million loss that would occur with a .9604 probability.) Expected shortfall for the combined portfolio is therefore (.04/2.5) x \$20 million + (2.46/2.5) x \$11 million = \$11.144 million. Characteristic of a subadditive risk measure, the expected shortfall in the combined portfolio of losing projects is less than the sum of the expected shortfall associated with each individual project:

$$\$11.144 \times 10^6 = ES_{p_{1+2}} \leq \$8.2 \times 10^6 + \$8.2 \times 10^6 = ES_{p_1} + ES_{p_2} \ . \qquad (16)$$

Although expected shortfall for any confidence interval is derived directly from *VaR* for that interval, only expected shortfall is subadditive and coherent. The reason for this apparent anomaly stems from the mathematical properties of the two measures:

> A risk measure can be characterized by the weights its assigns to quantiles of the loss distribution. *VaR* gives a 100% weight to the *X*th quantile and zero to other quantiles. Expected shortfall gives equal weight to all quantiles greater than the *X*th quantile and zero weight to all quantiles below the *X*th quantile. ... [A] spectral risk measure is coherent (*i.e.*, it satisfies the subadditivity

condition) if the weight assigned to the qth quantile of the loss distribution is a nondecreasing function of q. Expected shortfall satisfies this condition. *VaR* does not, because the weights assigned to quantiles greater than X are less than the weight assigned to the Xth quantile (Hull, 2012, p. 190).

4.2. Because expected shortfall is not an elicitable risk measure, it eludes meaningful backtesting

The virtues of subadditivity and coherence notwithstanding, expected shortfall does not represent a platonically ideal risk measure. Its principal problem is that it cannot be reliably backtested in the sense that forecasts of expected shortfall cannot be verified through comparison with historical observations. This is the primary respect in which *VaR* holds a regulatory advantage *vis-à-vis* expected shortfall as a measure of risk. *VaR* is easily backtested. Depending on the probability distribution by which market risk is modeled, the quantile at which *VaR* is measured not only identifies the frequency with which we would expect to encounter legally significant losses, but also sets the level of the loss that triggers regulatory attention. For instance, the one-day, 99 percent *VaR* drawn from a parametric analysis of an equity portfolio indexed to the S&P 500 and presumed to exhibit one-day standard deviation of 100 basis points can be expected lose at least \$23,260 on two or three days a year (1 percent of 252 trading days).

As a matter of intuition, this flaw is readily understood: expected shortfall purports to measure the full extent of risks in the tail of a loss distribution, including much deeper losses that are theoretically possible but are not observed during the relevant backtesting period. There is also a mathematical explanation of this weakness. In formal terms, expected shortfall is not *elicitable* (Gneiting, 2011). *VaR*, along with any other quantile-based risk measure, is elicitable (*ibid.*). The value of an elicitable risk measure is that it can be subjected to a consistent scoring function that properly reports the measure's reliability in forecasting future losses. Indeed, nearly the entire class of spectral risk measures, of which expected shortfall is a special case, is not elicitable. Whatever efforts we undertake to surmise the true shape and size of the tails of market-based loss distributions are just that, informed guesses in the face of incurable leptokurtic blindness.

The mathematically dictated failure of expected shortfall to satisfy the condition of elicitability puts the primary reform of Basel III in direct conflict with the primary reform of Basel 2.5. In endorsing *VaR*, Basel II embraced the primary risk measure that had taken root after the financial services industry's collective response to the global stock market collapse of 1987. Basel 2.5 addressed one of *VaR*'s known flaws: its fragility during periods of extreme economic stress, stemming from the failure of *VaR* to account for the shape and size of all risks in the tail beyond the quantile at which regulators direct banks to conduct *VaR* analysis. Stressed *VaR*, a crucial element of Basel 2.5, is designed to require banks to manage risk according to historic benchmarks that reflect the most severe threats to banks' trading books. Basel III's likely adoption of expected shortfall in place of *VaR* addresses another of the older, more established measure's known defects: *VaR*'s failure to satisfy subadditivity and coherence. But expected shortfall in turn eludes the backtesting enshrined in and directed by Basel 2.5. The newer, more sophisticated risk measure turns out to suffer from an intractable mathematical defect of its own.

The conflict between elicitable *VaR* and coherent expected shortfall has led critics of Basel III to decry what they regard as the Committee on Banking Supervision's quixotic quest for an unattainable mathematical ideal:

> [T]here is more to risk measurement than the choice of a *"risk measure"*: statistical robustness, and not only *"coherence,"* should be a concern for regulators and end users ... . [R]obust risk estimation ... requires the explicit inclusion of [a] statistical estimation step (Cont *et al*., 2010, p. 620).

5. Summary and conclusions

Uniquely among human endeavors, mathematics boasts *"a beauty cold and austere, ... without any appeal to any part of our weaker nature, without the gorgeous trappings of painting or music, yet sublimely pure, and capable of a stem perfection such as only the greatest art can show"* (Russell, 2008, p. 60). As one poet has expressed the sentiment: "Euclid alone has looked on Beauty bare" (Millay, 2003, p. 52).

Mathematics delivers an elegant denouement to the Basel accords' struggle to identify an ideal measure for market risk in global banks' trading books. It turns out that there is exactly one spectral risk measure that is both coherent and elicitable. Among spectral risk measures, only negative expected value satisfies both of these conditions (Ziegel, 2013). To return to *that* measure of risk would be to discard not only modern and post-modern portfolio theory, but also the entire edifice of quantitative finance grounded in on the notion that expected return must be weighed against volatility, beta, or some other measure of risk derived from the second moment of probability distributions. In other words, unless quantitative finance and the law are prepared to forswear reliance on the full dispersion of market results, the choice of *any* risk measure besides expected value, whether *VaR*, expected shortfall, or some other variation on quantile-based or spectral risk measurement, forces a choice between coherence and elicitability.

"Every year, if not every day, we have to wager our salvation upon some prophecy based upon imperfect knowledge" (United States Supreme Court, 1919, p. 630]. International

banking regulation under the Basel accords achieves precisely this much and nothing more — on the strength of quantitative risk measures specifying 99 percent confidence over holding periods of ten trading days.

Acknowledgements

This paper was presented on April 17, 2013, at Georgetown Law Center's colloquium on international financial regulation, and on October 12, 2013, at the 76th International Atlantic Economic Conference in Philadelphia. I appreciate comments by Christopher J. Brummer, Adam S. Candeub, Ephraim Clark, Daniel Martin Katz, José-María Montero Lorenzo, Jeffrey Sexton, Tino Sonora, Ben Walther, and Johanna F. Ziegel. Shawn Lowe Desai and Emily Strickler provided capable research assistance. Special thanks to Heather Elaine Worland Chen.

References

- Abbt, M., Limpert, E. and Stahel, W. (2001). Log-Normal Distributions Across the Sciences: Keys and Clues, *Bioscience,* **51**(5), pp. 341-352.

- Acerbi, C. and Tasche, D. (2002). On the Coherence of Expected Shortfall, *Journal of Banking and Finance*, **26**(7), pp. 1487-1503.

- Aitchison, J. and Brown, J. (1957). *The Lognormal Distribution, with Special Reference to its Uses in Econometrics*, Cambridge University Press, Cambridge, UK.

- Allen, L., Boudouk, J. and Saunders, A. (2004). *Understanding Market, Credit, and Operational Risk: The Value at Risk Approach*, Blackwell, Malden, MA.

- Artzner, P., Delbaen, F., Eber, J. and Heath, D. (1999). Coherent Measures of Risk, *Mathematical Finance*, **9**(3), pp. 203-228.

- Bangia, A., Diebold, F., Schuermann, T. and Stroughair, J. (1998). *Modeling Liquidity Risk, with Implications for Traditional Market Risk Measurement and Management,* WP[99-06], University of Pennsylvania, Philadelphia, PA. Available at: http://fic.wharton.upenn.edu/fic/papers/99/9906.pdf

- Bank for International Settlements (2011). *Interpretive Issues with Respect to the Revisions to the Market Risk Framework,* Basel, Switzerland. Available at: http://www.bis.org/publ/bcbs193a.pdf

- Basel Committee on Banking Supervision (2004). *International Convergence of Capital Measurement and Capital Standards: A Revised Framework*, Bank for International Settlements, Basel, Switzerland. Available at: http://www.bis.org/publ/bcbs107.pdf and http://www.bis.org/publ/bcbs107b.pdf

- Basel Committee on Banking Supervision (2011). *Messages from the Academic Literature on Risk Measurement for the Trading Book*, WP[19], Basel, Switzerland. Available at: http://www.bis.org/publ/bcbs_wp19.pdf

■ Basel Committee on Banking Supervision (2012). *Consultative Document: Fundamental Review of the Trading Book*, Basel, Switzerland. Available at: http://www.bis.org/publ/bcbs219.pdf

■ Benninga, S. and Wiener, Z. (1998). Value-at-Risk (*VaR*), *Mathematics in Education and Research*, **7**(4), pp. 39-45.

■ Bervas, A. (2006). Market Liquidity and Its Incorporation into Risk Management, *Financial Stability Review*, 8 (May), pp. 63-79. Available at: http://www.banquefrance.fr/fileadmin/user_upload/banque_de_france/publications/Revue_de_la_stabilite_financiere/etud2_0506.pdf

■ Brown, A. (2007). *On Stressing the Right Size*, GARP Risk Review, Jersey City, NJ.

■ Chen, J. (2012). Postmodern Disaster Theory, WP[11-17], Michigan State University College of Law, Michigan.

■ Cont, R., Deguest, R. and Scandolo, G. (2010). Robustness and Sensitivity Analysis of Risk Management Procedures, *Quantitative Finance*, **10**(6), pp. 593-606.

■ Daníelsson, J. and Zigrand, J. (2006). On Time Scaling of Risk and the Square-Root-of-Time Rule, *Journal of Banking and Finance*, **30**(10), p. 2702.

■ Dowd, K. (1998). *Beyond Value at Risk: The New Science of Risk Management*, Wiley, Chichester, UK.

■ Einhorn, D. (2008). *Private Profits and Socialized Risk*, GARP Risk Review, Jersey City, NJ.

■ European Commission (2014). *"Solvency II": Frequently Asked Questions*, Internal Markets & Services DG. Available at: http://ec.europa.eu/internal_market/insurance/docs/solvency/solvency2/faq_en.pdf

■ Farber, D. (2011). Uncertainty, *Georgetown Law Journal*, **99**(4), pp. 901-959.

■ Federal Deposit Insurance Corporation (1999). *Risk-Based Capital Standards: Market Risk*, Federal Register, **64**(2), p. 19.035.

■ Fisk, P. (1961). The Graduation of Income Distributions, *Econometrica*, **29**(2), pp. 171-185.

■ Forsey, H. (2001). The Mathematician's View: Modelling Uncertainty with the Three Parameter Lognormal, in Sortino, F., Satchell, S. (Eds.) *Managing Downside Risk in Financial Markets*, Butterworth-Heinemann, Oxford, UK, pp. 51-58.

■ Gneiting, T. (2011). Making and Evaluating Point Forecasts, *Journal of the American Statistical Association*, **106**(494), pp. 746-762.

■ Grau, W. (Ed.) (1999). *Five Guidelines on Market Risk: Stress Testing*, Österreichische Nationalbank, Vienna, Austria.

■ Huber, P. (1981). *Robust Statistics*, John Wiley & Sons, Hoboken, New Jersey.

■ Hull, J. (2012). *Risk Management and Financial Institutions*, Pearson, Upper Saddle River, NJ.

■ Jorion, P. (2006). *Value at Risk: The New Benchmark for Managing Financial Risk*, McGraw-Hill, New York.

■ Macchiarola, M. (2009). Beware of Risk Everywhere: An Important Lesson from the Current Credit Crisis, *Hastings Business Law Journal*, **5**(2), pp. 294-297.

■ McNeil, A., Frey, R. and Embrechts, P. (2005). *Quantitative Risk Management*, Princeton University Press, Princeton, NJ.

- Millay, E. (1922). Euclid Alone Has Looked on Beauty Bare, in McClatchy, J. (Ed.) *Selected Poems*, Harper Perennial, New York.
- Mina, J. and Xiao, J. (2001). *Return to RiskMetrics: The Evolution of a Standard.* Available at: http://www.wu.ac.at/executiveeducation/institutes/banking/sbwl/lvs_ws/vk4/rrmfinal.pdf
- Nocera, J. (2009). Risk Mismanagement, *New York Times Magazine*, January 4, 2009 edition, p. MM24.
- Nordhaus, W. (2011). The Economics of Tail Events with an Application to Climate Change, *Review of Environmental Economics and Policy*, **5**(2), pp. 242-43.
- Páfka, S. and Kondor, I. (2001). Evaluating the RiskMetrics Methodology in Measuring Volatility and Value-at-Risk in Financial Markets, *Physica A*, **299**(1-2) p. 309.
- Portnoy, S. and He, X. (2000). A Robust Journey in the New Millennium, *Journal of the American Statistical Association*, **95**(452), p. 1331.
- Russell, B. (1988). The Study of Mathematics, in Allen, G. (Ed.) *Mysticism and Logic, and other Essays*, Spokesman, Nottingham, UK, pp. 58-73.
- Sheppard, K. (2013). *Financial Econometrics Notes*, University of Oxford, UK. Available at: http://www.kevinsheppard.com/images/ b/bb/Financial_Econometrics_2013-2014.pdf
- Taleb, M. (2009). *Report on The Risks of Financial Modeling, VaR and the Economic Breakdown*, Available at: http://gop.science.house.gov/Media/hearings/oversight09/sept10/taleb.pdf.
- United States Supreme Court (1919). *Abrams v. United States*, United States Reports, **250**(316), pp. 616-629.
- Whitehead, C. (2011). Destructive Coordination, *Cornell Law Review*, **96**(2), p. 341.
- Zangari, P. (1996). *RiskMetrics Technical Document*, J. P. Morgan, New York.
- Ziegel, J. (2013). Coherence and Elicitability, *Mathematical Finance* (forthcoming 2014). Available at: http://arxiv.org/pdf/1303.1690v2