risks

MDPI

Article

Aggregation of Incidence and Intensity Risk Variables to Achieve Reconciliation

Clive Hunt [1,*] **and Ross Taplin** [2,*]

[1] Private Practice, Perth, WA 6009, Australia
[2] School of Accounting, Curtin Business School, Curtin University, Bentley, WA 6102, Australia
* Correspondence: clivehunt@bigpond.com (C.H.); R.Taplin@curtin.edu.au (R.T.)

Received: 22 August 2019; Accepted: 21 October 2019; Published: 25 October 2019

Abstract: The aggregation of individual risks into total risk using a weighting variable multiplied by two ratio variables representing incidence and intensity is an important task for risk professionals. For example, expected loss (EL) of a loan is the product of exposure at default (EAD), probability of default (PD), and loss given default (LGD) of the loan. Simple weighted (by EAD) means of PD and LGD are intuitive summaries however they do not satisfy a reconciliation property whereby their product with the total EAD equals the sum of the individual expected losses. This makes their interpretation problematic, especially when trying to ascertain whether changes in EAD, PD, or LGD are responsible for a change in EL. We propose means for PD and LGD that have the property of reconciling at the aggregate level. Properties of the new means are explored, including how changes in EL can be attributed to changes in EAD, PD, and LGD. Other applications such as insurance where the incidence ratio is utilization rate (UR) and the intensity ratio is an average benefit (AB) are discussed and the generalization to products of more than two ratio variables provided.

Keywords: mean; Basel Accord; IFRS 9; credit risk; insurance risk; attribution; probability of default; loss given default

1. Introduction

Credit risk modelling for capital and provisioning in banking is an example of the modelling context where this paper is relevant (other common examples are considered in Section 2). The Basel Accord is a regulatory framework for capital held in case of a severe economic downturn (Basel Committee on Banking Supervision 2006) and the International Financial Reporting Standard 9—Financial Statements (IFRS 9) is a framework for provisioning for unpaid debt (International Accounting Standards Board 2014). Under these regulatory frameworks (henceforth referred to as the Basel Accord and IFRS 9 respectively), a bank's lending book is made up of exposures, for each of which a probability of default (PD) and a loss given default (LGD) is predicted. Each exposure also has a monetary size called the exposure at default (EAD), and the LGD represents the proportion of EAD lost. A standard calculation defines the expected monetary loss (EL) for this exposure as

$$\mathrm{EL} = \mathrm{EAD} \times \mathrm{PD} \times \mathrm{LGD} \quad (1)$$

Summing EL for all exposures then estimates the total expected monetary losses for the whole lending book.

EL calculation is illustrated in Table 1 with a simple example consisting of three exposures at two points in time (period 1 and period 2). These three exposures might be three individual loans or summaries of three portfolios, such as home loans, credit cards, and personal loans. For each period each exposure has a value for EAD, PD, and LGD[1]. For example, in period 1 exposure 1 has an EAD of \$115,000, a PD of 0.5% and an LGD of 90%. This paper takes these values as given, their calculation arising from risk models following frameworks such as the Basel Accord and IFRS 9. From Equation (1), the product of these three values gives an EL of \$518. The total EL of \$1998 (aggregate row of Table 1) for period 1 is the sum of the EL values for each exposure and is of great importance to the business, investors, auditors and regulators (for example, in the context of provisioning it forms an assessment of liabilities on the balance sheet and feeds directly into the profit and loss statement of the company).

Table 1. Lack of reconciliation for credit risk data.

Exposure	Period 1				Period 2			
	EAD	**PD**	**LGD**	**EL**	**EAD**	**PD**	**LGD**	**EL**
1	\$115,000	0.50%	90%	\$518	\$115,000	1.00%	90%	\$1035
2	\$25,000	5.00%	80%	\$1000	\$25,000	5.00%	80%	\$1000
3	\$160,000	3.00%	10%	\$480	\$160,000	2.50%	10%	\$400
Aggregate [1]	\$300,000	2.21%	46.5%	\$1998	\$300,000	2.13%	46.5%	\$2435
Product [2]	\$300,000 × 2.21% × 46.5% = \$3081				\$300,000 × 2.13% × 46.5% = \$2976			

[1] Aggregates are: sum for exposure at default (EAD) and expected loss (EL), EAD-weighted arithmetic means for probability of default (PD) and loss given default (LGD). [2] Products of Aggregates are EAD × PD × LGD.

Defining the aggregate EL as the sum of the EL for each exposure is most natural and it is equally natural to also define the aggregate EAD as the sum of the EAD for each exposure (Table 1). Appropriate definitions for aggregates of PD and LGD are less obvious, but a measure of central tendency (such as a mean) provides a more useful contender than a sum. We use the generic term "aggregate" for this practice of deriving a single figure (sum or mean) to represent the summary of a set of values. Exposure-weighted means $\overline{\text{PD}}$ and $\overline{\text{LGD}}$ (aggregate row of Table 1) are natural choices and are indeed specified in at least one Basel-related standard, reporting form ARF 113.1 (Australian Government 2018). For example, the aggregate for PD in period 1 of 2.21% is the weighted average of 0.5%, 5%, and 3% (with weights of 115,000, 25,000, and 160,000). These provide useful summaries of the PD and LGD risk components of the exposures within a portfolio but have the following undesirable properties.

First, the EAD weighted means for PD and LGD do not have a reconciliation property defined by Equation (1). That is, for period 1, the product of \$300,000 × 2.21% × 46.5 = \$3081 is far removed from the true value for EL of \$1998. This brings into question the value of using these EAD weighted means as summary aggregates for PD and LGD. EAD weighted means of PD and LGD may, of course, be useful for other purposes, especially when they are interpreted in isolation, but they are not so useful when they are interpreted in the context of a system of variables used to define the key quantity of interest, EL.

Second, Table 1 shows an increase in EL from \$1998 in period 1 to \$2435 in period 2. An obvious question is what caused this increase: to what extent is the increase in EL due to an increase in portfolio size (EAD), an increase in the risk of default (PD) or an increase in the inability of defaulted loans to be repaid (LGD)? An analyst attempting this attribution analysis using the aggregates in Table 1 would have the illogical situation of trying to explain how the increase in aggregate EL is due to a decrease in aggregate PD (from 2.21% to 2.13%) while aggregate EAD and LGD remain unchanged. Thus, changes

[1] For now we consider these values to have come from models designed to estimate the respective quantities (prospective analysis) but they could also come from observed data (historical analysis) when monitoring observed exposures at default, default rates and losses given default.

in EAD weighted means for PD and LGD are not useful, and potentially misleading, in the context of explaining changes in EL.

Third, both the aggregates for PD and LGD are treated in isolation: for example, the aggregate for LGD is logically weighted by EAD however PD is ignored. For period 1 (in Table 1), a case can be made that the aggregate for LGD should depend more heavily on exposure 3 than exposure 1 because the PD is much higher for exposure 3 than exposure 1. In Section 3 we discuss this point in detail and argue that aggregates for PD and LGD should be considered jointly using a system-approach based on Equation (1).

This brings us to the nub of our paper, which is how to aggregate PD and LGD. In particular, we desire the aggregates for PD and LGD, along with the aggregates (sums) of EAD and EL, to satisfy Equation (1) when these aggregates are used instead of values for individual exposures (we call this the reconciliation property). That is, we desire aggregates of PD and LGD that are useful in the context of understanding EL, including why EL changes over time or why EL differs between segments of a portfolio. This paper introduces novel "joint-ratio means" as aggregates that satisfy this reconciliation property and explores other properties, including how to use these aggregates to perform attribution analysis (where changes in EL are attributed to changes in EAD, PD, and LGD). Understanding EL in this way is important not only to analysts requiring appropriate methods to conduct their work, but also management within the business, auditors, regulators, investors and policy makers.

The remainder of this paper is organized as follows. Section 2 elaborates on the context for these joint ratios, with examples from banking and insurance, and reviews the relevant literature. Section 3 discusses alternative weighted means and Section 4 introduces the joint-ratio means. Section 5 shows how changes in EL can be attributed to EAD, PD, and LGD. Section 6 considers how results can vary depending on the aggregation path followed. Discussion is presented in Section 7 followed by concluding comments. The discussion so far has exactly two ratio variables, PD and LGD, but Appendix A shows how the joint-ratio idea, and its reconciliation property, generalizes naturally to m ratio variables.

2. Literature Review

The study of means and their properties has a long history. For example, Kolmogorov (1930) provided an axiomatic treatment, which is described in Muliere and Parmigiani (1993) and also succinctly in de Carvalho (2016). Familiar means include arithmetic-, geometric-, harmonic- and power—means, mode, as well as those depending on ordering of the data, such as median and trimmed means[2]. When weights are available, appropriately weighted versions of the foregoing can also exist. Muliere and Parmigiani (1993), and notably the references they list, track contributions on means across a century by many leading thinkers in statistics and decision theory. However, this literature addresses means in a univariate, compartmentalized setting. This paper, by contrast, takes a system approach, where means for all variables in the system are considered simultaneously to produce desirable properties for the system rather than considering each variable in isolation.

The purpose is an important motivation in general and this has been recognized for means in particular, as explained by Chisini (1929), quoted here as translated by Muliere and Parmigiani (1993, p. 423):

> The search for a mean has the purpose of simplifying a given question, by substituting to many values a single summary value, and leaving the overall picture of the problem under consideration unchanged [. . .] One should not be thinking about the mean of two or more

2 We do not consider robust versions such as trimmed means and medians further because these can be misleading by hiding extreme observations (such as high LGD values close to 1 amongst the majority of values close to 0). Furthermore, means are useful because when multiplied by size they give a total (for example, the average PD multiplied by the size of the portfolio gives a total of defaults). Note that means exist and are useful irrespective of the distribution of PD, EAD and LGD.

values, but only about the mean of those values with reference to the evaluation of a quantity that depends on them.

Falk et al. (2005) provide an accessible illustration of this viewpoint by showing four different ways to calculate mean vehicle speed, depending on the purpose. Our paper follows this doctrine by considering the purpose of calculating aggregate EL when determining appropriate aggregates (means) for PD and LGD.

The common context addressed in this paper is the modelling of portfolio risk as an aggregation of individual risks, indexed by $\{i|i = 1, \ldots, n\}$, modelled at this granular level via a product of quantities

$$\mathrm{EL}_i = W_i r_{1i} r_{2i} \tag{2}$$

which are then summed to arrive at the aggregate (total) risk:

$$\mathrm{EL} = \sum_{i=1}^{n} \mathrm{EL}_i = \sum_{i=1}^{n} W_i r_{1i} r_{2i} \tag{3}$$

The quantities W_i, r_{1i} and r_{2i} are, for now, all required to be strictly positive, but otherwise, have no constraint (Appendix A discusses how this may be relaxed). Note subscripts in Equation (2) make it explicit how Equation (1) is applied at a granular level (indexed by i), which are summed in Equation (3) to obtain aggregate (total) risk. While there may be no formal requirement for Equation (1) to hold at the aggregate level as well, we argue this is a useful property in practice because Equation (1) defines the system in which these components exist.

In the banking example, W_i, r_{1i}, and r_{2i} represent EAD, PD, and LGD respectively, with EL_i the quantum of expected loss. Table 2 summarizes this and other contexts where this formulation arises but note symmetry exists between ratio 1 and ratio 2 and it makes no difference whether, for example, PD and LGD are swapped. With reference to the literature, we now expand upon the other examples in Table 2.

Table 2. Example contexts for the basic model.

Context	Weight (W_i)	Ratio #1 (r_{1i})	Ratio #2 (r_{2i})
Credit risk—historic	EAD	Default rate	LGD
Credit risk—prospective	EAD	PD	LGD
Insurance—generic	Exposure value	Incidence	Intensity
Insurance—Private Health	Number of members	Utilisation rate (UR)	Average benefit (AB)

One distinction highlighted in the first two rows of Table 2 is whether calculations are performed retrospectively (using observed data) or prospectively (using model predictions). In retrospective contexts, incidence is the rate at which risk events occurred for that group of exposures (the partition will usually be groups of exposures rather than individual exposures since, when dealing with historical facts, individual exposures could only return a binary rate of 1 or 0 depending on whether the risk event occurred or not). In prospective settings, the incidence would usually be a probability of occurrence of a risk event within some time window such as the following 12 months. These probabilities may result from some model and may be derived at a very granular level such as individual bank loans, or at some level of grouping (pools). Intensity is a variable representing the severity of the risk event, such as the fraction of the exposure that is lost, and as with incidence it could either reflect actual experience (retrospective) or be a prospective model.

This aggregation issue also occurs in health insurance (see for example Practice Guidance 699 issued by the Institute of Actuaries of Australia (IAA 2012, sct. 8.3.2c)) where (see Table 2) the insurance paid equals the number of members (weighting variable) multiplied by the product of two ratio variables: an incidence ratio given by the utilization rate (UR) and an intensity ratio given by the average benefit (AB). In this setting, the product of the two ratios UR $\times$ AB is called the "drawing rate".

As a related aside, it can be said that the context of this paper arises when the drawing rate is modelled as a product of two (or more) risk components—such as UR × AB in the health example, or PD × LGD in credit risk.

The Basel Accord literature makes extensive recommendations for the modelling of credit risk using the quantities EAD, PD and LGD. The focus of the Basel Accord is on risk-based capital calculations, but banks also use similar formulations in contexts such as provisioning models (for example, IFRS 9), where the modelled expected loss is instead called "collective provisions". The literature includes criticism of the Basel Accord calculations, including the specification of values for asset correlation when calculating unexpected loss (Gabbi and Pietro 2013). The asset correlation captures how PD are correlated over time in the sense that the default of exposures can be linked to a common factor representing economic stress. PD and LGD can also be correlated over time, in the sense that both PD and LGD can be linked to this same factor. This paper acknowledges these issues however they do not directly impact the nub of this paper. Given the international reach of the Basel Accord for credit risk in banks, henceforth this paper uses this context, noting that this nomenclature also accords with the common risk classification system for the actuarial profession proposed by Kelliher et al. (2013). The natural translation for other contexts is left to the reader.

Other approaches are possible, such as Figini and Giudici (2013), which assume risk is measured on an ordinal scale such as high, medium and low. This is common in other areas such as operational risk. However, our paper assumes risk can be measured on a ratio scale. This is common in many applications and indeed a requirement for credit risk modelling under Basel Accord and IFRS 9.

Whilst the joint-ratio means that we introduce in Section 4 appear to be a new type of mean, it is more helpful to understand them as a new setting for means—one that takes into account the entire system in which all variables exist and interact and where the means satisfy the reconciliation property, all flowing from Equation (1). This setting does not appear to have been explicitly addressed in the literature (a comprehensive recent review of aggregation theory can be found in the two-part overview by Grabisch et al. (2011a, 2011b)). Our approach has been to present this novel idea in the community most likely to find it applicable in practice, namely the Risk community, exemplified by the credit risk context. Some further clarifications of the modelling context are now noted.

The subscript $\{i|i = 1, \ldots, n\}$ can represent granular risks at many levels. These include accounts, loans, customers, and portfolios. For example, the three rows of Table 1 might represent the home loan, credit card and personal loan portfolios of a retail bank (n = 3) with Equation (3) defining the EL of the bank. In prospective credit risk modelling, where models predict future losses, it is feasible that every account will be individually modelled and hence n could be in the millions. By contrast, historic default rates are likely to be available as aggregates of groupings (dimensions) such as product, cohort, demography, or acquisition channel—perhaps only a handful in number although more if taken in combination. For our purposes, the size of the partition is not important. What is however important is what intermediate stages may have been involved in summarizing upwards from the most granular level. Section 5 addresses this topic under the heading of "aggregation path invariance", our name for a subject addressed in the literature for the univariate context under various names, such as "associativity" (the fourth axiom in Kolmogorov 1930), or "strong idempotency" (Calvo et al. 2004). This invariance property requires the same final result even if aggregation is carried out via any intermediate stages: for example, if means are calculated for different partitions of the data and those means, duly weighted, are then aggregated into a final mean.

3. Aggregates (Means) for PD and LGD

Table 3 expands on the example in Table 1 by calculating a variety of weighted means for PD and LGD. Aggregates (weighted means) for PD are calculated using weights of EAD (mean = 2.21% and 2.13% for period 1 and period 2) and with weights of EAD × LGD (mean = 1.43% and 1.75%). For example, the mean of 1.43% is the weighted mean of 0.5%, 5.0% and 3.0% with weights of 103,500, 20,000 and 16,000 respectively. Similarly, aggregates for LGD are calculated using weights of EAD (46.5% for both periods) and EAD × PD (30.2% and 38.0%).

Table 3. Weighted Means of Table 1 exposures and their reconciliation to the actual EL.

Aggregation Type	Period 1			Period 2		
	PD	LGD	EL [1]	PD	LGD	EL [1]
Mean (weighted by EAD)	2.21%	46.5%	$3081	2.13%	46.5%	$2976
Mean (weighted by EAD × other ratio)	1.43%	30.2%	$1295	1.75%	38.0%	$1992
Joint-ratio mean	1.78%	37.4%	$1998	1.93%	42.1%	$2435

[1] EL values equal the product of the aggregate EAD, aggregate (mean) PD and aggregate (mean) LGD. The actual EL is $1998 (period 1) and $2435 (period 2).

Either weighting (by EAD or by EAD × another ratio) can be justified as useful in specific contexts. For example, the LGD mean of 46.5% (EAD weighting) is useful as a measure of central tendency for the LGD within the portfolio and can be calculated independently of PD values. The LGD mean of 30.2% (weighted by EAD × PD) also takes into account the probability of default. This is arguably a more sensible summary of the portfolio LGD since the LGD is less relevant (requiring a lower weight) when the risk of default is lower. For example, consider an exposure with a PD of 0 (or immaterially different to 0). The LGD of this exposure should be ignored (or immaterial) in the calculation of the mean LGD as the exposure cannot default, so its actual LGD value is irrelevant. The argument is analogous to the desire to weight by EAD: an exposure with zero EAD (or immaterial EAD) should receive no (or negligible) weight in the calculation of the mean LGD since it will have negligible impact on EL. Hence the argument to weight by EAD suggests the mean LGD should also be weighted by PD.

Similarly, for PD the portfolio mean can be EAD weighted or EAD × LGD weighted. The latter is justified on the basis that exposure with LGD negligibly different to 0 has negligible impact on EL (this exposure could arguably be removed before calculation of an EAD weighted mean). While the EAD weighted mean PD might measure the proportion of the exposure expected to default, the mean PD weighted by EAD × LGD is a more relevant summary from the perspective of expected loss (and arguably a business perspective) because in this context exposure with negligible LGD is irrelevant regardless of its PD.

The above discussion also applies when calculations are performed retrospectively (using historical data), although business practice may differ. In this case, the default outcome is known for each exposure, so PD takes the value of 0 (not default) or 1 (default). In this retrospective calculation using observed data, analysts might naturally weight by observed default as well as EAD (equivalently, weight by EAD after removing non-defaulted exposures). In this retrospective application, the decision to weight by PD amounts to the decision to condition on defaults (as the name loss given default implies). Importantly, this discussion highlights how both weighted means are useful and meaningful depending on the context.

For prospective calculation, when aggregates are required for model prediction (rather than for actual outcomes), the distinction or even existence of these two alternative means might be less well recognized. For example, when summarizing the LGD of exposures within a portfolio (for example to compare two portfolios), weighting by EAD may be more natural. Guidelines such as the Basel Accord consider PD and LGD separately and this might encourage PD and LGD to be aggregated independently: in practice these functions might even be performed by different teams or at different times (the authors do not recommend this business practice of considering components of the system in isolation, they recommend a system-approach that considers the context of the entire system).

Mean PD weighted by EAD × LGD may seem less intuitive for several reasons. First, there is a natural time ordering of events: the first PD to determine defaults and then LGD to determine loss from those defaults. While this conceptualization of events in time may be useful elsewhere, it is not required or necessarily useful in the calculation of EL. Second, under retrospective calculations, the EAD × LGD weighting for mean PD would require known losses. In practice, this can be time consuming to calculate or even unavailable as the default remains unresolved (that is, the exposure has not emerged from default, so the loss amount is not available yet).

Neither weighting by just EAD or by the product of EAD and the other ratio variable produces values that reconcile in the aggregate to EL. Table 3 shows a value of \$3081 for \$300,000 × PD × LGD in period 1 when EAD weighted aggregates are used and \$1295 when weights of EAD × other ratios are used (the correct aggregate EL is \$1998). Thus, despite the above discussion of advantages and disadvantages of these weightings, using either consistently throughout does not produce aggregates that reconcile to the true EL.

This lack of reconciliation can be resolved by using different versions of the weightings for PD and LGD. Note that in period 1 using the EAD weighted PD as the aggregate for PD (2.21%) and the (EAD × PD) weighted mean of LGD as the aggregate for LGD (30.2%) produces the correct EL when applied in Equation (1) as \$300,000 × 2.21% × 30.2% equals \$1998. Similarly, \$300,000 × 1.43% × 46.5% also equals \$1998. It is trivial to show this equality always hold by substituting definitions of these weighted means into Equation (1) and cancelling terms: as long as one ratio is aggregated using weights of EAD and the other ratio is aggregated using weights of EAD × another ratio then the aggregates reconcile to the correct EL.

We refer to these asymmetric solutions as Sequential Weighted Arithmetic Means (SWAM) noting that the calculation is performed in sequence: first one ratio is aggregated using weights of EAD only and then the other ratio is aggregated using weights equal to the product of EAD and the other ratio. Taking PD as the first ratio to be weighted by EAD may appear more natural but this is probably only due to a conceptualization that PD precedes LGD: this has no mathematical basis. We expand on deficiencies of the non-symmetrical SWAM approach in Section 7 but first, we provide the more elegant mathematical solution of joint-ratio means.

4. Joint-Ratio Means

Consider a set of n triples (indexed $i = 1, \ldots, n$), comprising a weight variable W_i and two ratio variables r_{1i} and r_{2i}:

$$\{(W_i,\ r_{1i}, r_{2i}) \big| i = 1, \ldots, n\} \tag{4}$$

For example, the data for period 1 in Table 1 has three triples (exposures) comprising the weight variable EAD and the ratio variables PD and LGD.

The joint-ratio aggregate is defined to be the triple

$$\left(\sum W_i, \overline{r}_1^*, \overline{r}_2^*\right) \tag{5}$$

where

$$\overline{r}_1^* = \sqrt{\frac{(\sum W_i r_{1i})(\sum W_i r_{1i} r_{2i})}{(\sum W_i)(\sum W_i r_{2i})}} \tag{6}$$

and

$$\overline{r}_2^* = \sqrt{\frac{(\sum W_i r_{2i})(\sum W_i r_{1i} r_{2i})}{(\sum W_i)(\sum W_i r_{1i})}} \tag{7}$$

and all summations are over all triples ($i = 1, \ldots, n$). We refer to the triple $(\sum W_i, \overline{r}_1^*, \overline{r}_2^*)$ as the joint-ratio aggregate (since $\sum W_i$ is a sum) and $\overline{r}_1^*$ and $\overline{r}_2^*$ as joint-ratio means (joint-ratio to emphasize these single values depend on both ratio variables). Equations (6) and (7) have an intuitively useful interpretation as the geometric mean of two weighted means: one weighted by the weight variable W_i and the other weighted by (W_i × another ratio). This interpretation also reveals why we can consider $\overline{r}_j^*$ ($j = 1$ or 2) to be a mean of the raw values $\left\{r_{ji} \big| i = 1, \ldots, n\right\}$ of that jth ratio, hence justifying the nomenclature 'joint-ratio *mean*'.

By substituting the definitions given by Equations (6) and (7) the following equation holds:

$$\left(\sum W_i\right)\overline{r}_1^*\overline{r}_2^* = \sum(W_i r_{1i} r_{2i}) \tag{8}$$

Equation (8) expresses the fact that the aggregates from Equation (5), when used in Equation (1), reconcile to the correct risk aggregate found by summing results from Equation (1) applied to the *n* triples given by Equation (4). That is, joint-ratio means reconcile to the correct EL.

Table 3 summarizes these results for the example in Table 1. The joint-ratio mean for PD in period 1 is 1.78%, which can be calculated directly from Equation (6) or as the square-root of 2.21% × 1.43% (the geometric mean of the EAD weighed and EAD × LGD weighted means of the raw PD values). Similarly, the joint-ratio mean for LGD is $\sqrt{46.5\% \times 30.2\%}$ = 37.4%. Furthermore, \$300,000 × 1.78% × 37.4% equals the correct EL of \$1998 so the joint-ratio aggregates reconcile.

These joint-ratio means have the advantage that no priority is given to either ratio. As illustrated in Section 3, a case can be made for either EAD weighted or EAD × LGD weighted means for the mean PD. These arguments apply equally to LGD. However, regardless of which weighting is preferred using the same weightings for PD and LGD fails the reconciliation property. The symmetric joint-ratio mean is a compromise between these two possibilities that applies the same logic to each ratio rather than giving one ratio preferential treatment in terms of receiving the preferred weighting (whichever weighting is considered preferred). Both ratios, PD and LGD, are considered equal (or interchangeable), just as they are in Equation (1) defining EL.

Appendix A contains generalizations of the symmetric joint-ratio means to the case of 3 or more ratio variables.

5. Attributing Change in EL to Changes in the Individual Components

A common analytical or monitoring request is to explain the difference between two EL values, in terms of the individual components EAD, PD, and LGD. This arises most naturally when the EL of a portfolio at two different points in time are being compared. For example, a trend analysis might compare EL values for consecutive quarters or financial years, with a Chief Risk Officer wanting to know why estimated collective provisioning has increased (see Table 1). Another example might be comparing two different portfolios (or two segments of one portfolio) at the same point in time.

A little algebra shows the difference between two EL values can be decomposed into four components as follows:

$$\begin{aligned}\Delta\text{EL} \quad &= \ \text{EL}_2 - \text{EL}_1 \\ &= \ \Delta\text{EAD} \times \overline{\text{PD}} \times \overline{\text{LGD}} \text{ (change in EL due to a change in EAD)} \\ &+ \overline{\text{EAD}} \times \Delta\text{PD} \times \overline{\text{LGD}} \text{ (change in EL due to a change in PD)} \\ &+ \overline{\text{EAD}} \times \overline{\text{PD}} \times \Delta\text{LGD} \text{ (change in EL due to a change in LGD)} \\ &+ (\Delta\text{EAD} \times \Delta\text{PD} \times \Delta\text{LGD})/4 \text{ (residual)}\end{aligned} \tag{9}$$

where all terms refer to aggregates, $\Delta\text{EAD} = \text{EAD}_2 - \text{EAD}_1$, $\overline{\text{EAD}} = (\text{EAD}_1 + \text{EAD}_2)/2$ and likewise for the other components, with subscripts indicating the points in time (or portfolios) being compared.

Equation (9) can be interpreted as breaking down ΔEL, the change in EL, into the amounts of change attributable to the change in each component. For example, the change in EAD contributed $\Delta\text{EAD} \times \overline{\text{PD}} \times \overline{\text{LGD}}$ (in absolute \$ amount) to the change in EL. The attribution is not arithmetically exact without the residual term $(\Delta\text{EAD} \times \Delta\text{PD} \times \Delta\text{LGD})/4$ in Equation (9). Due to the product of several typically small changes in the three components, this is usually negligible in magnitude and can safely be ignored (or split between the three primary terms).

Using the joint-ratio means in Table 3, the increase in EL of \$437 from period 1 to period 2 can be attributed to PD in the amount \$300,000 × (1.93% − 1.78%) × (42.1% + 37.4%)/2 = \$180.65 and similarly \$256.85 to LGD (as EAD does not change, \$0 is attributed to EAD and the residual equals \$0).

Although the LGD has not changed from period 1 to period 2 for any of the exposures, the increase in PD for exposure 1 (with high LGD) and the decrease in PD for exposure 3 (with low LGD) increases the likelihood of exposure with high LGD defaulting and therefore from an EL perspective the LGD of the portfolio has increased.

In contrast, under SWAM, if the EAD weighted mean is used for PD and the (EAD × PD) weighted mean is used for LGD, the change in EL of $437 from period 1 to period 2 can be attributed as $300,000 × (2.13% − 2.21%) × (38.0% + 30.2%)/2 = −$76.72 to PD and $300,000 × (2.13% + 2.21%)/2 × (38.0% − 30.2%) = $514.22 to LGD. Thus, a negative EL is attributed to PD because the EAD weighted mean PD decreased even though the PD has increased where it matters more (exposure 1 with the high LGD). Compared to the analysis based on joint-ratio aggregates (JRA), under SWAM a larger increase in EL must be attributed to LGD because of the negative attribution to PD. So far, this example shows that either JRA or SWAM can be used to make attribution calculations since either method satisfies the reconciliation property (Equation (8)). We claim that the more useful interpretation in this example is the one arising from JRA, the nub being that JRA considers PD to have increased, whereas SWAM reflects a decrease. Given that the only raw figures in Table 1 that changed between periods 1 and 2 were PD figures, and that the consequent EL increased, it is not coherent to base an interpretation on a decrease in PD. Rather, from perusal of Table 1 one is forced to focus on how PD changes have correlated with values of the other ratio (LGD). This theme of a 'joint' or 'system' approach lies at the heart of the JRA and we return to it in Section 7.

Finally, if all exposures have one variable (EAD, PD, or LGD) changed by multiplying by a constant (so all exposures have the same relative increase or decrease) then sensibly all of the change in EL is attributed to the variable that is changed. However, it is naïve to expect, and not generally the case, that changing only one variable will result in all the change in EL being attributed to that variable. In Table 3, only PD values are changed but some of the resulting change in EL is attributed to LGD because the relative importance of the LGD values changes as a consequence. The situation is analogous to when the EAD for one exposure is changed: this is likely to result in the EAD weighted PD (and LGD) changing and hence sensibly not all the change in EL being attributed to EAD.

6. Lack of Aggregation Path Invariance

Joint-ratio means, along with many other summary statistics such as medians and trimmed means, lack a property which we call "aggregation path invariance", or invariance for short. Aggregation path invariance means the same value is obtained for aggregates regardless of the order of any partial steps taken during the aggregation process. This property is important if aggregates are required not only for the whole portfolio but also for segments of the portfolio. For example, if aggregates are required for several products (credit cards, home loans, and personal loans) then invariance requires the aggregation of all accounts in the portfolio (in one step) to equal the aggregation of the aggregates for these segments. We consider aggregation path invariance in depth in this section because it is the main disadvantage of joint ratio means compared to simple means and therefore requires consideration when adopting their use.

For example, if exposures 2 and 3 in period 1 of Table 1 are aggregated then the resulting joint-ratio aggregates (Equation (5)) are ($185,000, 3.67%, 21.8%). If this result is then aggregated with exposure 1 (period 1) in Table 1, then the resulting aggregates are ($300,000, 1.85%, 36.1%). While similar, these differ from the aggregates of 1.78% for PD and 37.4% for LGD when all three exposures are aggregated in one step (Table 3). Hence the joint-ratio means for PD and LGD differ depending on the path taken to perform the aggregation. The weight variable (here EAD) does satisfy aggregation path invariance because of the associativity property of addition as it applies to the set $\{W_i | i = 1, \ldots, n\}$. No matter how many stages or what partitions are chosen, one will always arrive at the correct aggregate EAD of the whole portfolio.

Importantly, aggregation path invariance also holds for the EL. This holds in general, similarly to the weight variable, by associativity of addition as it applies to the set $\{W_i r_{1i} r_{2i} | i = 1, \ldots, n\}$, together with the reconciliation Equation (8). No matter how many stages or what partitions are chosen, recursive application of Equation (8) guarantees that, whatever subset X of the exposures is currently under consideration, the product of the three components of the joint ratio aggregates for X will equal $\sum_{i \in X} W_i r_{1i} r_{2i}$.

The positive message above is that using joint-ratio means as an aggregation method, there will always be perfect consistency in the EAD and EL figures no matter how varied the analyses involved. The negative message is that the PD and LGD aggregate figures will usually vary between analyses having the same end-point if the starting points and/or aggregation pathways were different. Hence a risk officer requesting, say, the summary of the Home Loans portfolio, will see different figures for the aggregate PD and LGD depending on whether the aggregation is from the individual exposures, or is derived from a summary of Home Loans by geographic zone (but the EAD and EL will agree exactly).

The extent to which the aggregate PD and LGD can vary depending on aggregation path involves only one degree of freedom, since the product PD × LGD is invariant, being bound to equal the invariant quotient EL/EAD. Hence any upward variation in PD is matched by an inverse movement in LGD, as seen in the examples above where the two step aggregation gives a drawing rate (aggregate PD × aggregate LGD) of 1.85% × 36.1% = 0.67% which is equal to the corresponding product 1.78% × 37.4% for the one step aggregation (Table 3), with both equal to the EL/EAD ratio of \$1998/\$300,000.

7. Discussion

A focus on analytics arises naturally from the requirement to monitor the individual components of risk, as extensively provided for by the Basel Accord for credit risk in the banking industry. The Basel Accord prescribes the monitoring of PD and LGD individually, this is appropriate because separate models will have been used to predict each of these components. Similarly, in the Health Insurance context the Institute of Actuaries of Australia (IAA 2012, para. 8.3.2c) draws attention to monitoring UR and AB components separately.

This shows that the industries are interested not only in analytics in general but specifically at the level of the individual risk components such as PD and LGD (or UR and AB). Joint-ratio means squarely target this need within the framework of the overall system of estimation and their use is recommended for this purpose.

In the Basel Accord (and IFRS 9) formulation of EL in Equation (1) applied at the granular level, aggregates for EAD and EL as summations of individual values are intuitive. Aggregates of PD and LGD as an overall summary of their individual values require more careful thought. This paper fills this need by considering alternative aggregates, from simple weighted means to joint ratio means, and exploring their properties within the EL system. In particular, joint ratio means are shown to have an important and useful reconciliation property so that EL = EAD × PD × LGD holds for the aggregates as well as at the individual level. This avoids apparent logical inconsistencies when these aggregates are interpreted within the system. For example, changes in aggregate EL can be more easily attributed to changes in the aggregate EAD, PD, and LGD components.

The importance of this system-wide approach cannot be under-estimated. Not only is it similar to the desire to ensure the default definition for PD and LGD are equivalent, but business management, modellers, auditors and regulators need to ensure the system as a whole is functioning as intended, not just the components separately. The attribution of changes in aggregate EL to changes in aggregate EAD, aggregate PD, and aggregate LGD is just one illustration of this system wide approach to examining aggregates (or means).

The Sequential Weighted Arithmetic Mean (SWAM) approach also results in aggregates that reconcile. The decision on which aggregates are most suitable will depend on the context in which these aggregates are to be used however when making this decision there are several reasons why the Joint-Ratio Aggregate (JRA) approach might be considered preferable.

First, the product of PD and LGD in Equation (1) for EL is symmetric: interchanging PD and LGD gives the same result since multiplication is commutative. This symmetry is observed by the JRA formula but not by SWAM. Second, this lack of symmetry could be resolved by taking an average of the two SWAM results (depending on whether PD or LGD is considered first with an EAD weighted average) but this is precisely what the JRA does by taking the geometric mean of these two SWAM approaches.

Third, there are two defendable SWAM approaches and it is debatable which is more appropriate. While the PD weighted by EAD might be considered more natural based on the fact that default comes first and then loss follows, the reverse weighting can also be considered natural. When considering predictions for a set of loans (rather than historical observed results) it is a valid thought process to invert the sequence of events: first, multiply EAD × LGD and consider that as the "bank funds exposed to risk", with PD quantifying the probabilities of each of those loss outcomes. Furthermore, any argument that the aggregate for PD should be weighted by EAD but not by LGD ignores the clearly higher risk (at least from an expected loss perspective) of exposure with a higher LGD. An argument to ignore LGD in the calculation of aggregate PD might be applied to EAD as well, suggesting unweighted means of PD are appropriate but these are clearly far removed from a system-wide approach taking into account expected losses.

In the private health insurance (PHI) context (Table 2), our experience is that UR and AB are treated symmetrically by analysts, certainly to a greater extent than generally applies to PD and LGD in the banking context. The essential difference here is that defaults are rare, sporadic and individualised by comparison with utilisation rates of a fixed schedule of medical services. Whilst it is trivially true that a service must happen before its benefit amount can be measured, it is not unusual for the business to consider causality in the opposite direction, such as when planning changes to the benefits schedule and contemplating how that might affect utilisation patterns. The precedence of UR or AB is not clear and we propose that a symmetric treatment is appropriate in this case.

The above reasons can become stronger, or more important because they are less obvious when more sophisticated analysis is performed such as the attribution analysis in Section 5. Furthermore, a case can be made for presenting multiple results including EAD weighted means, means weighed by the product of EAD and the other ratio variable, as well as the JRA for both PD and LGD. Doing so allows the reader to not only note whether the results differ depending on the type of mean calculated but also use whichever mean is most suited for a particular purpose. For example, if a new PD scorecard is being proposed and the result on EL of just changing PD is desired by the business, then the EAD × LGD weighted mean for PD is of greatest interest. However, if the change in EL from the previous reporting period is reported to the market then an attribution analysis using other aggregates, such as the joint-ratio aggregates, may also be of interest. This paper encourages intellectual consideration of these factors to bring analytics to a higher level rather than a simplistic reporting of EAD weighted means.

Attribution analysis should not, of course, be limited to examining how changes in EL might be due to changes in EAD, PD, and LGD. Analysis should also consider the effect of individual cases since means can be heavily influenced by a few outliers. For example, a change in EL might be due to a change in one account, the accounts of one customer, or the accounts in one segment of the portfolio.

When making comparisons between different points in time such as last year versus this year, the lack of invariance can be minimized by following the same aggregation pathway in each instance. For example, if the figures for last year were first aggregated into the state, and then to the grand total, apply the same aggregation stages to the current year. This makes it more likely that a consequent difference in mean PD truly reflects a difference in individual PD values rather than an artifact of aggregation pathway. This is already common practice in other contexts: for example decisions concerning whether multiple accounts on a loan (for example, a variable interest home loan and a fixed interest home loan secured by the same property, or two credit cards owned by the same customer) should be first aggregated into a single observation or treated as two separate loans. Path aggregation issues are also minimised when similar exposures are aggregated first (such as when loans are aggregated together within their portfolios of credit card or home loans).

Finally, although the argument in this paper for the joint ratio aggregates is self-contained and will be of immediate use to institutions managing credit risk, there are several avenues for future research. These include incorporating these ideas into the development of PD and LGD credit risk models, extending the ideas from expected loss to unexpected loss (including considerations of asset correlation, assessing the impact of distributions for PD, EAD, and LGD (and in particular the influence of extreme values), and calculating standard errors and confidence intervals to assess the statistical significance of changes in quantities such as the joint ratio means (and issues of asset correlation).

8. Conclusions

A risk model context that is common in banking and insurance is identified and exemplified by the widely used Basel Accord formulation of EL = EAD × PD × LGD applied at the granular level. Aggregation of EAD and EL is straightforward, by summation, but the calculation of a mean value for PD and LGD is less clear-cut. This paper demonstrates problems with using EAD weighted means for both PD and LGD and proposes a new type of mean for this situation that has the property of reconciling in the aggregate.

Regardless of the aggregates chosen for PD and LGD, this paper explores the properties of alternative aggregates and argues that a system approach is desirable that not only considers the estimation of these aggregates jointly but does so within the EL framework in which these quantities exist. We see considerable prospects for the joint-ratio means to add value to risk analytics.

Author Contributions: Conceptualization of the joint ratio mean, C.H.; all other aspects including methodology, software, validation, formal analysis, investigation, and writing—review and editing, C.H. and R.T.

Funding: This research received no external funding.

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

Appendix A

This appendix gives an algebraic presentation of the joint-ratio aggregate concept, generalized to any order $m > 2$, where m is the number of ratio variables. We describe the three-ratio case before proceeding to the general case.

Given a set of n tuples, comprising a weight variable W_i and three ratio variables r_{1i}, r_{2i} and r_{3i}:

$$\{(W_i,\ r_{1i}, r_{2i}, r_{3i})\big| i = 1, \ldots, n\}$$

the joint-ratio aggregate tuple is defined to be the tuple.

$$\left(\sum W_i, \overline{r}_1^*, \overline{r}_2^*, \overline{r}_3^*\right)$$

where:

$$\overline{r}_1^* = \sqrt[3]{\frac{\sum W_i r_{1i}}{\sum W_i} \sqrt[2]{\frac{\sum W_i r_{1i} r_{2i}}{\sum W_i r_{2i}} \frac{\sum W_i r_{1i} r_{3i}}{\sum W_i r_{3i}}} \frac{\sum W_i r_{1i} r_{2i} r_{3i}}{\sum W_i r_{2i} r_{3i}}}.$$

$$\overline{r}_2^* = \sqrt[3]{\frac{\sum W_i r_{2i}}{\sum W_i} \sqrt[2]{\frac{\sum W_i r_{1i} r_{2i}}{\sum W_i r_{1i}} \frac{\sum W_i r_{2i} r_{3i}}{\sum W_i r_{3i}}} \frac{\sum W_i r_{1i} r_{2i} r_{3i}}{\sum W_i r_{1i} r_{3i}}}$$

and

$$\overline{r}_3^* = \sqrt[3]{\frac{\sum W_i r_{3i}}{\sum W_i} \sqrt[2]{\frac{\sum W_i r_{1i} r_{3i}}{\sum W_i r_{1i}} \frac{\sum W_i r_{2i} r_{3i}}{\sum W_i r_{2i}}} \frac{\sum W_i r_{1i} r_{2i} r_{3i}}{\sum W_i r_{1i} r_{2i}}}$$

Noting that most of the sum-products (such as $\sum W_i r_{1i}$ and $\sum W_i r_{1i} r_{2i}$) cancel, it can be confirmed that the following required reconciliation property holds:

$$\left(\sum W_i\right)\overline{r}_1^*\overline{r}_2^*\overline{r}_3^* = \sum (W_i r_{1i} r_{2i} r_{3i})$$

The remainder of this Appendix A explains the pattern for any order m, generalising the pattern that can be seen developing in the order 3 case above.

The data: There are n rows ("$(m+1)$-tuples") of data.

$$\{(W_i,\ r_{1i}, r_{2i}, \ldots r_{mi})\big| i = 1, \ldots, n\} \tag{A1}$$

each tuple comprising a weight variable $W_i > 0$ and m ratio variables r_{1i}, r_{2i}, ... , r_{mi}, all likewise required to be > 0.

Unless otherwise specified, summation is understood to be over all the tuples $\{i|i = 1, \ldots, n\}$.

The joint-ratio aggregate of the data in Equation (A1) is defined as the $(m+1)$-tuple:

$$\left(\sum W_i, \overline{r}_1^*, \overline{r}_2^*, \ldots, \overline{r}_m^*\right). \tag{A2}$$

where the definition of $\overline{r}_j^*$ is described in three stages below. In brief, $\overline{r}_j^*$ is constructed as the geometric mean of the m geometric means of all possible k-suborder weighted arithmetic means ($k = 0, \ldots, m-1$) of the set $\left\{r_{ji}\middle| i = 1, \ldots, n\right\}$. By this construction, $\overline{r}_j^*$ is in a strong sense a mean of the jth ratio variable.

The first stage of defining $\overline{r}_j^*$ is to form every possible k-suborder weighted arithmetic mean (WAM) of the set $\left\{r_{ji}\middle| i = 1, \ldots, n\right\}$ as follows:

$$WAM_j^{k,p} = \left(\sum r_{ji} W_i r_{p_1 i} r_{p_2 i} \ldots r_{p_k i}\right) / \left(\sum W_i r_{p_1 i} r_{p_2 i} \ldots r_{p_k i}\right). \tag{A3}$$

Sub-order $k = 0, 1, \ldots, (m-1\ )$ indicates how many of the other ratios (i.e., other than j) are used in the weighting scheme along with W_i. We do not allow any ratio variable to be used in weights against itself, hence p is taken to range over all $\begin{pmatrix} m-1 \\ k \end{pmatrix}$ choices $p = \{p_1, p_2, \ldots, p_k\} \subset \{1, 2, \ldots, m\} - \{j\}$.The second stage of defining $\overline{r}_j^*$ is to form, for each k, the geometric mean (GM) of all k-suborder WAMs:

$$GM_j^k = GM\left\{WAM_j^{k,p}\middle| over\ p\right\} k = 0, 1, \ldots, (m-1). \tag{A4}$$

Finally, $\overline{r}_j^*$ is defined as the geometric mean of the set of instances over k of Equation (A4):

$$\overline{r}_j^* = GM\left\{GM_j^k\middle| k = 0, 1, \ldots, (m-1)\right\}. \tag{A5}$$

The following desired reconciliation property can be shown to hold in this general case (proof omitted):

$$\left(\sum W_i\right)\prod_{j=1}^{m} \overline{r}_j^* = \sum W_i r_{1i} r_{2i} \ldots r_{mi}. \tag{A6}$$

Effectively, the joint ratio aggregate is a new kind of aggregation function, where the novelty stems from requiring the reconciliation property (A6) to hold.

References

Australian Government. 2018. Reporting Standard ARS 113.1 Advanced Internal Ratings-based (AIRB) Approach to Credit Risk. Available online: https://www.legislation.gov.au/Details/F2018L00394 (accessed on 22 October 2019).

Basel Committee on Banking Supervision. 2006. *Basel II: International Convergence of Capital Measurement and Capital Standards, A Revised Framework—Comprehensive Version*. Basel: Bank for International Settlements, Available online: https://www.bis.org/publ/bcbs128.htm (accessed on 4 February 2018).

Calvo, Tomasa, Radko Mesiar, and Ronald R. Yager. 2004. Quantitative weights and aggregation. *IEEE Transactions on Fuzzy Systems* 12: 62–69. [CrossRef]

Chisini, O. 1929. Sul concetto di media. *Periodico di Mathematiche* 9: 106–16.

de Carvalho, Michel. 2016. Mean, what do you mean? *The American Statistician* 70: 270. [CrossRef]

Falk, Ruma, Avital Lann, and Shmuel Zamir. 2005. Average speed bumps: Four perspectives on averaging speeds. *Chance* 18: 25–32. [CrossRef]

Figini, Silvia, and Paolo Giudici. 2013. Measuring risk with ordinal variables. *Journal of Operational Risk* 8: 35–43. [CrossRef]

Gabbi, Giampaolo, and Vozzella Pietro. 2013. Asset correlations and bank capital adequacy. *The European Journal of Finance* 19: 55–74. [CrossRef]

Grabisch, Michel, Jean-Luc Marichal, Radko Mesiar, and Endre Pap. 2011a. Aggregation functions: Means. *Information Sciences* 181: 1–22. [CrossRef]

Grabisch, Michel, Jean-Luc Marichal, Radko Mesiar, and Endre Pap. 2011b. Aggregation functions: Construction methods, conjunctive, disjunctive and mixed classes. *Information Sciences* 181: 23–43. [CrossRef]

IAA. 2012. Practice Guideline 699.01 Pricing and Financial Projections for Private Health Insurers. IFRS 9—Financial Instruments. Available online: https://www.actuaries.asn.au/library/Standards/HealthInsurance/2012/PG699.01-final-September2012.pdf (accessed on 11 October 2017).

International Accounting Standards Board. 2014. IFRS 9—Financial Instruments. Available online: http://www.aasb.gov.au/admin/file/content105/c9/AASB9_12-14.pdf (accessed on 4 February 2018).

Kelliher, Patrick O. J., D. Wilmot, J. Vij, and Paul J. M. Klumpes. 2013. A common risk classification system for the actuarial profession. *British Actuarial Journal* 18: 91–121. [CrossRef]

Kolmogorov, Andreĭ Nikolaevich. 1930. Sur la notion de la moyenne. *Atti Accad. Naz. Lincei* 9: 388–91.

Muliere, Pietro, and Giovanni Parmigiani. 1993. Utility and means in the 1930s. *Statistical Science* 8: 421–32. [CrossRef]