# Analysis of financial events under an assumption of complexity

## Yifei Li

*Sydney Business School, University of Wollongong, Wollongong, New South Wales, Australia*

## John Evans*

*Centre for Analysis of Complex Financial System, Sydney, New South Wales, Australia*

## Abstract

The financial system can be shown to be a complex adaptive system consisting primarily of a federation of systems and systems of systems. There are significant similarities between the characteristics of natural systems and financial systems suggesting that the type of analysis employed in understanding natural systems could have application in financial system analysis. Cladistics analysis has been used extensively for analysis of biological systems and has accordingly been used in the social sciences for some years but a rigorous justification for adopting the analysis has not been undertaken. This paper discusses the appropriateness of applying cladistics analysis to financial systems, and then considers the appropriate methodology to be adopted for analysis of different financial events.

## 1. Introduction

Traditional statistical analysis remains the dominant methodology for analysing financial systems, and by way of illustration, in the 2017 editions of a top-rated financial journal, of the 62 papers published, 55 included traditional statistical methods. But the financial system can be shown to be a complex adaptive system (CAS) which means the system is continually changing and the outcome of the system's reactions to internal or external disruptions cannot be predicted from a reductionist analysis. This emergence phenomenon of the financial system has been recognised for some time with consequent questioning of the appropriate tools for analysis. We would argue that traditional statistical approaches have limited application for modelling the ambiguity risks in the financial systems as defined by Ganegoda & Evans (2014). Danielsson (2008) and Danielsson *et al*. (2016) both observed that traditional statistical models were fragile as to their assumptions and mechanisms for understanding financial systems, as well as inconsistent due to the endogenous uncertainty inherent in the financial systems. It seems that understanding financial systems under such complexity requires

*Correspondence to: John Evans, Centre for Analysis of Complex Financial System, PO Box 363 Summer Hill, Sydney, New South Wales, Australia. E-mail: jevans@pgeaus.onmicrosoft.com

360

a solution that can provide a holistic picture of the system. We are of the view that cladistics analysis is capable of providing such a holistic view of the relationships occurring in the financial systems. Cladistics analysis is a method to classify events in a CAS into related sets. This method has been applied to different areas, and whilst initially developed for biological applications, it has now been extended to the social sciences. In the social sciences, the methodology has been used to analyse management systems (Mitleton-Kelly, 2003), product development (McCarthy *et al.*, 2006), cultural inheritance in social systems (Matthews *et al.*, 2013), operational risk in banks (Li *et al.*, 2017*a*), world economic risks (Evans *et al.*, 2017) and credit risks (Shi *et al.*, 2018*a*). In these applications outside of biology, there has been an assumption that cladistics analysis was appropriate as the systems being analysed were CAS. Whilst a system being a CAS is essential to justify the adoption of cladistics analysis, other criteria apply and also the exact methodology to be applied will vary depending on the particular characteristics of the CAS. In this paper we will discuss the justification for adopting cladistics analysis for financial events, the different algorithms that are applicable to particular sets of financial events, and provide some insights into the value of cladistics analysis of financial events.

## 2. Similarities of Biological and Financial Systems

The most important similarity between biological and financial systems is that they are both CAS. CAS and their attributes have been widely discussed by John Henry Holland (1995), Arthur, Durlauf & Lane (1997), Cilliers (1998), Mitleton-Kelly (2003), Mitchell (2009) and Holland (2014). Whilst there is not a commonly accepted definition of a CAS, it is generally accepted there are four basic attributes to identify a CAS, namely, there are numerous components, there is no central control, there are interactions amongst the components, and there is emergence of the system as a result of interactions. The major financial institutions were shown by Schweitzer *et al.* (2009) to be highly connected with loops existing between the financial institutions, indicating that financial institutions are strongly inter-connected. Schweitzer *et al.* (2009) argued that such interdependence may result in instability of the network which is a signal of emergence of a system. The financial industry also has no central control as was recognised as early as Adam Smith (1776) when he talked about "an invisible hand" operating in the markets. Similarly, in the financial industry, there is no global control for the trading activities of participants, and regulators are concerned only with specific geographic areas as shown by Evans & Li (2018*a*) who demonstrated that the global banking system was a federation of systems (FOS)[1]. It is relatively easy to observe the emergent property of a CAS in the financial industry as it is impossible to predict the market change by observing one or two financial institutions. The interactions of all the financial institutions create the emergent property with the activities of agents in the financial industry being influenced by other agents, resulting in coevolution (ul-Haq, 2005; Song & Thakor, 2010). Song & Thakor (2010) found that co-evolution in banks was generated by the effect of including securitisation of other banks' assets in bank equity capital. Evans & Li (2018*a*) argued that the extent of the interdependence of the global banking system was so high as to require a change in the regulation of banks from an FOS to an SOS basis. The above discussion leads to the conclusion that the global financial system presents the essential characteristics of a CAS, that is, numerous agents, interactions among agents, no central control and emergence. Allan *et al.* (2010) demonstrated the parallels of evolution in financial risks and biology in that financial risks have unique characteristics similar to DNA in biology. Allan *et al.* (2012)

---

[1] Whilst globally the financial system is an FOS, within counties the financial system may well be a system of systems (SOS) which vary from FOS in the extent to which there is some central control.

investigated financial risk evolution using Darwinian criteria and found financial risk evolution satisfied all the criteria, namely variation, competition, inheritance, accumulation of modifications and adaptation. The conceptual parallels between biological evolution and financial events can be summarised as:

1. Characteristics: in biological evolution, the characteristics are phenotype, that is, there are observable characteristics and molecular sequence changes can affect the phenotype (Griffiths *et al.*, 2005). For financial events, the characteristics are determined from the descriptions and records of the events and the characteristics are an abstraction and summarisation of these descriptions rather than the records themselves.
2. Evidence: the evidence of biological evolution includes observation of fossils and current species. The evidence of evolution of financial events is similarly based on historical records and descriptions of the characteristics of events and descriptions of the characteristics of current events.
3. Random variation and selection: in biological evolution, the variation is caused by some environmental determinants or happenstance, hence selection occurs from natural selection or genetic drift (Lande, 1976). For financial events the environment (e.g. innovation, regulation and transactions) is the main source of variation with risk management or controls (or the lack of controls) being the mechanism that determines what new combinations of characteristics will emerge.
4. Inheritance: Inheritance exists for financial events through the occurrence of events with the same combinations of characteristics as historical events.

Given the justification of financial systems as a CAS and these conceptual parallels of biological evolution and emergence of financial events, we can draw the conclusion that it is feasible to apply cladistics analysis in studies related to financial events.

## 3. Cladistics Analysis

In applying cladistics analysis, it is important to consider the format of the data, as this will affect the algorithm selection, the encoding methods that transform the data to meet the needs of the particular analysis, and the interpretation of the resultant phylogenetic tree. The investigation of evolution in biology has used two different data types, namely, molecular data and morphological data. Morphological data records the form, structure and structural features of species, including appearance (e.g. colour) and internal structure (e.g. bones). Molecular data analyses DNA and proteins to gain information on evolutionary relationships. There is debate as to whether to adopt molecular-based analysis or morphological-based analysis, but Wiens (2000) pointed out the most common cause of incongruence was due to under sampling of characters and taxa. Hillis (1995) argued that a combination of molecular and morphological data would yield a better estimation of the true evolution. Wiens (2000) argued that the data used could be molecular or morphological, or a combination of these two, so as long as the selection can be justified and the results are properly interpreted. Encoding in cladistics analysis refers to the transformation of the characteristics into a format that can be used by the various algorithms available. Pleijel (1995) delineated four different encoding methods for cladistics analysis which essentially relate to whether or not the presence of a characteristic needs to be encoded as well as the absence of the characteristic. The different encoding methods are:
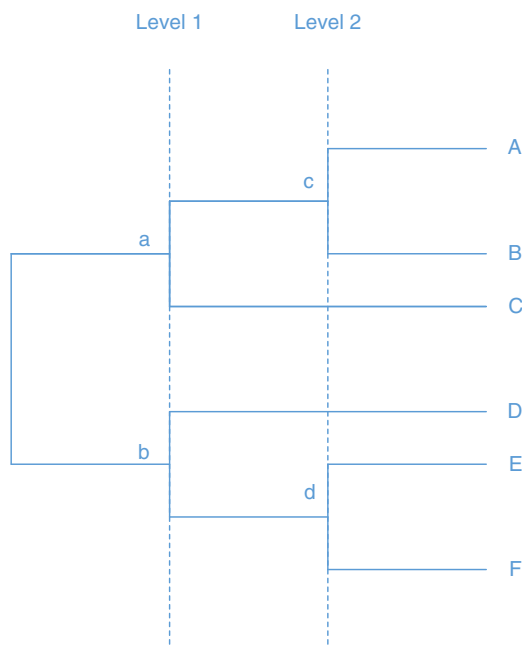
362

**Figure 1.** An example of a tree.

1. Linked multi-states which requires the absence and the presence for each characteristic to be encoded but there are fixed characteristic combinations permissible;
2. Independent multi-states which requires the absence or presence of each characteristic to be encoded, but there are no fixed combinations of characteristics;
3. Independent binary states which only requires either the absence or the presence of the characteristic to be encoded but where all characteristics need to be encoded;
4. Binary states for all characteristics where only the presence or absence of characteristic is encoded and not all characteristics are required to exist.

A classic way to illustrate the outcome of a cladistics analysis is by estimating phylogenetic trees. In biology, a phylogenetic tree (or cladogram) is a graph that presents the inferred evolutionary relationships among different species. A phylogenetic tree can be transformed into various shapes (e.g. diagonal-up, rectangular-right, rectangular-up, diagonal-down and circle, as discussed by Baum & Smith (2013)), but these shapes are just different ways of showing the same inferred evolution. The rectangular-right tree is visually easy to understand and an example is illustrated in Figure 1. A phylogenetic tree diagram consists of leaves, branches and nodes. The leaves, for example, A, B, C and D in Figure 1, represent different species (organisms, genes) in an evolutionary context. A rotation of branches under a node will not change the relationships. For instance, denoting the leaves under node a in Figure 1 as ((A, B), C) if it is rotated as ((B, A), C) or (C, (A, B)), it is still the same tree.

There are two different approaches for inferring phylogenetic trees, namely, distance-based methods and character-state based. Distance-based methods usually construct a phylogenetic tree based on a

distance matrix of pairwise genetic distances (Felsenstein, 1988). The main distance-based methods include cluster analysis such as UPGMA (unweighted pair group method using arithmetic averages, Sokal & Michener (1958)) and WPGMA (weighted-pair group method with arithmetic means, Sokal & Michener (1958)) which assume a consistent evolutionary rate, minimum evolution (Kidd & Sgaramella-Zonta, 1971; Rzhetsky & Nei, 1993) and minimises the total distance and neighbour-joining (Saitou & Nei, 1987; Studier & Keppler, 1988). Character-state-based approaches, or sequence-based methods rely on the state of the character, and all possible trees are evaluated to generate the one that optimises the evolution. The main character-state-based methods include maximum likelihood methods (Felsenstein, 1981), which includes Bayesian methods, and parsimony methods (Camin & Sokal, 1965; Kluge & Farris, 1969; Fitch, 1971). Bayesian methods and maximum likelihood methods are both statistical inference methods. The parameters of Bayesian methods are variables with distributions whilst the parameters of maximum likelihood methods are unknown constants. Bayesian inference relies on prior probabilities (Rannala & Yang, 1996; Yang & Rannala, 1997). Both maximum parsimony and maximum likelihood methods are character-based methods, and they rely on different phylogenetic characteristics, for example, genetic, morphological and molecular attributes to construct the trees. Character-state-based methods are often considered more powerful than distance-based methods (Rastogi, Mendiratta, & Rastogi, 2008), as they use raw data while distance-based methods transform raw data into a distance matrix which introduces information loss. These two methods of constructing the trees are based on different philosophies. The main assumption of parsimony is simplicity (Farris, 1983), which results in the minimum number of homoplasies (i.e. a character that different species share is not inherited from a common ancestor). Farris (1983) first made the justification that the minimisation of ad hoc hypothesis of homoplasy maximises the explanatory power. Sober (1975) considered simplicity as a matter of how much extra information has to be obtained to enable the theory to answer the research question, while the less information is needed, the more informative is the theory. Some others (Queiroz & Poe, 2001; Kluge, 2006; Wiley & Lieberman, 2011) attribute parsimony to "Ockham's razor," which states that simpler hypotheses are preferred over complex ones. The alternative methodology, maximum likelihood, is a statistical concept based on the probability of given data. There is debate around the philosophy of phylogenetic inference (Popper, 1959; Popper, 1983; Popper, 2002; Helfenbein & DeSalle, 2005). Popper changed the term "degree of conformation" to "degree of corroboration" (Popper, 2002) and argued that hypotheses should survive from the most severe tests. Popper (2002) saw truth as eternal while corroboration as temporal and further gave a formula for the degree of corroboration:

$$C(h, e, b) = \frac{p(e, hb) - p(e, b)}{p(e, hb) - p(eh, b) + p(e, b)}$$

There are four terms to be examined: probability ($p$), background knowledge ($b$), empirical evidence ($e$) and hypothesis ($h$). The logical probability "of a statement is complementary to its degree of falsifiability" (Popper, 2002), and "the support given by $e$ to $h$ becomes significant only when $p(e, hb) - p(e, b) \gg 1/2$" (Popper, 1983). For hypothesis with high content, $p(eh, b)$ is close to 0 (K. Popper, 2002). For a given ($e$) and ($b$), $p(e, b)$ is constant and therefore the trees with highest $p(e, hb)$ will be the tree with the strongest corroboration. In the context of cladistics analysis, background knowledge ($b$) is the assumptions inherent in the method, empirical evidence ($e$) is the data, and hypothesis ($h$) is the hypothesis of relationships (Queiroz, 2004). Carpenter (1992), Siddall & Kluge (1997), Carpenter, Goloboff, & Farris (1998) and Farris, Kluge, & Carpenter (2001) argued that the parsimony methodology corresponds to the philosophy of Karl Popper, and it is a better method compared to the principle of likelihood, whilst Queiroz (1988), Faith & Cranston (1991), Queiroz & Poe (2001), Queiroz (2004) and Queiroz (2014) doubted the argument and argued the equally

364

compatible, or even superiority of likelihood with Popperian corroboration. Farris (1973), Felsenstein (1973) and Goldman (1990) argued that parsimony can be expressed under a likelihood framework because C($h$, $e$, $b$) is straightforward in a likelihood framework. As well as the philosophical debate, there has also been debate at the practical level of the alternative methodologies. Wright & Hillis (2014) and Puttick *et al.* (2017) claimed that Bayesian analysis outperforms parsimony methods for discrete morphological data but O'Reilly *et al.* (2016) wrote "only minor differences are seen in the accuracy of phylogenetic topology reconstruction between the Bayesian implementation … and parsimony methods". Schrago, Aguiar, & Mello (2018) used empirical data to compare Bayesian inference and maximum parsimony, and found more trees at the 95% confidence level for Bayesian inference compared to a maximum parsimony method and concluded Bayesian inference was less precise than maximum parsimony. Brown *et al.* (2017) and Goloboff *et al.* (2018) also recommended caution for the model being applied to morphological data as the methods applied by Puttick *et al.* (2017) and O'Reilly *et al.* (2017) might, as pointed out by Goloboff *et al.* (2018), cause long branch attraction for parsimony methods (Felsenstein, 1978). Hence, in practice, parsimony methods are at least not worse than, maximum likelihood methods (Puttick *et al.*, 2017) and it is the most widely applied method for morphological data (Wright & Hillis, 2014; Puttick *et al.*, 2017). Another consideration is what is known as long branch attraction. When the evolutionary rate is extremely unbalanced, there will be a long branch attraction which leads to inconsistency of tree estimation (Felsenstein, 1978). With two possible character states, each with a possibility of $P$ and $Q$ to change, when $P^2 \leq Q(1-Q)$, there might be a long branch attraction problem. For small $Q$, the situation is approximated by $P \leq \sqrt{Q}$ (Felsenstein, 1978). If there is a significant difference in the evolution rate, modified parsimony methods that reduce the impact of the evolutionary rate (Lake, 1987; Willson, 1999) or maximum likelihood methods, which are less sensitive to long branch attraction, should be considered. The final consideration is to recognise the necessity to align the research method and the research objective. There are several algorithms for constructing a cladistics tree using the parsimony criterion. Camin & Sokal (1965) introduced the first algorithm to apply parsimony in constructing a cladogram. Later, Kluge & Farris (1969) presented the Wagner parsimony algorithm for constructing a cladogram and generating the most parsimonious tree. Fitch (1971) and Farris (1973) introduced other methods for tree construction. These algorithms have different assumptions, with the main differences being:

1. Camin-Sokal parsimony assumes evolution is irreversible, that is, a derived character state cannot return to its ancestral state.

2. Wagner parsimony assumes evolution is reversible, and the rates of change in either direction is roughly the same. It also assumes ordered characters, that is, a change from state 3 to state 1 must pass through state 2.

3. Fitch parsimony assumes evolution is reversible with approximately the same change rate in each direction, and it considers all characters as unordered, that is, a change from state 3 to state 1 does not have to go through state 2.

4. Dollo parsimony (Dollo, 1893) assumes the transition from the ancestral state is very rare, but there is no restriction on transitions from derived state to ancestral state.

## 4. Applying Cladistics Analysis to the Financial System

The application of cladistics analysis to the financial systems is then not straightforward and it is necessary to ensure the research method and the research objective are aligned. One of the issues

that arises is that for biological molecular data there are only four nucleobases, namely adenine (A), cytosine (C), guanine (G), thymine (T) as the fundamental genetic code for DNA, and A, G, C, uracil (U) for RNA but it is difficult to identify comparable basic characteristics for financial events. The financial event characteristics are usually derived from descriptions of the events and are then estimations of underlying basic characteristics and are not constant across different financial events in the same way that the nucleobase characteristics are constant in biology. For instance, the bank operational risk studies (Li *et al.*, 2017*a*, 2017*b*) estimated the major drivers of operational risk events across different regions by transforming the descriptions of the events to some characteristics but other financial risk studies (Evans *et al.*, 2017; Shi *et al.*, 2018*a*) used very different characteristics to that of Li *et al.* (2017*a*, 2017*b*) as the research objectives were to analyse WEF global risks and credit risk. Morphological data which considers characteristics would then seem a better choice for analysis of financial events and since likelihood methods (as well as Bayesian methods) are model based, which is hard to generate for financial events, parsimony is a reasonable basis to adopt (Goloboff *et al.*, 2017). For financial event characteristic encoding, there are four main objectives:

1. The encoding should reflect the attributes under investigation. It is important to note that, unlike the application in biology, the encoding of financial events is not an objective process, and the selection of characteristics and the identification of different states must be chosen so as to correctly represent the information contained in the source data and match the purpose of the study.

2. The encoding should help reduce the total number of events. As the financial data may have millions of events, it is vital to reduce the number of events to a practical level. One way to limit the number of unique events is to apply binary encoding to the source data.

3. The encoding should reflect the underlying assumption in the algorithm being applied, for example, if the Camin-Sokal algorithm is applied, then the cause of events should be set constantly as state 1.

4. Continuous characteristics should be transformed into discrete data for cladistics analysis. There are several methods to transform continuous data into discrete data, for example, simple gap-coding (Mickevich & Johnson, 1976), segment coding (Colless, 1980) and generalised gap-coding (Archie, 1985). These methods create gaps to produce discrete codes for continuous data (Kitching *et al.*, 1998).

Based on these objectives, we would argue that for financial events, encoding method 3 as outlined in section 3 would be appropriate as using binary encoding for the presence and absence of a characteristic results in categorical data, which reduces the states (compared to encoding methods 1 and 2) and the number of characteristics (compared to encoding method 4). However, this encoding method, as discussed before, will require careful selection of the characteristics to present the attributes of the financial events. Another issue worth mentioning is the information loss when transforming continuous characteristics. Wiens (2001) proposed a method for transforming continuous characteristics based on gap-weighting (Thiele, 1993), which leads to less information loss than gap-coding (Mickevich & Johnson, 1976). The interpretation of the phylogenetic trees for financial events is also quite different to that for biological application which focuses more on the structure of the leaves (A, B, C, D, E, F) in Figure 1, to provide classification of species but in financial event studies, they represent the financial events that occurred. Nodes, for example, a, b, c and d in Figure 1, correspond to lineage-splitting events, and in financial event studies, they are the

366

characteristics of events. The branches, that is, the connections between nodes, also have a different meaning for financial event studies where these branches specify relationships rather than an evolutionary path, as there is no time line involved and financial event analysis is an unrooted tree. Given the hierarchical structure of the phylogenetic trees for financial events the left most characteristics can be referred to as level 1 characteristics, for example, nodes a and b in Figure 1. The second characteristics along the path are denoted as level 2 characteristics, for example, nodes c and d in Figure 1, and so on. The level 1 characteristics for financial events can be thought of as the most systemic characteristics as they apply to the most events and therefore are of most interest in controlling the occurrence of the events as controlling these characteristics will have the most impact on financial losses. There are also other issues to consider, including the rate of change of the combination of characteristics that result in a new financial event as this will not be a constant as demonstrated in Li *et al.* (2018) where the rate of change for credit defaults and capital markets was shown to significantly increase as a tipping point was approached but operational risk events did not show a similar rate of change over the same time period. In determining the appropriate algorithm for financial event studies, since Dollo parsimony includes the unrealistic assumption for financial events that transition from an ancestral state is very rare and Fitch parsimony is a generalised Wagner methodology, we recommend either Fitch parsimony or Camin-Sokal parsimony. If the number of characteristics is small and all the states are irreversible, Camin-Sokal parsimony is recommended, as it allows a simple and intuitive way to transform data. In the bank operational risk studies, Li *et al.* (2017*b*) used Camin-Sokal method to construct the trees. However, if the characteristics are in multiple states, or the number of events and characteristics are large, Fitch parsimony will be more efficient.

## 5. Empirical Illustrations

To illustrate the value of cladistics analysis relative to traditional statistical analysis we have included a brief comparative analysis for credit risks, operational risk events and motor vehicle insurance claims.

## 5.1. Credit risks

Ali *et al.* (2016) used a multi-factor regression methodology to assess the relevance of characteristics of bankrupt individuals and Table 1 shows the statistical results of their analysis.

The statistically significant causes were identified as age, couple (i.e. couples are more likely to go bankrupt than single people), metropolitan, (i.e. people living in cities are more likely to go bankrupt than people living outside of cities), clerical/machinery (i.e. people engaged in clerical jobs or jobs associated with machinery are more likely to go bankrupt) and income. Shi *et al.* (2018*b*) used cladistics analysis on data from the same source as Ali *et al.* (2016) and was able to provide much richer insights into bankruptcies using the Carmin-Sokal algorithm. The consistent systemic characteristics identified by the cladistics analysis were age, gross income, spouse income, no real assets and major city, which shows differences to the results for Ali *et al.* (2016) and suggests that bankruptcies occur predominately within the pre-retirement population and importantly, are driven by what Shi *et al.* (2018*b*) defined as socio-economic issues rather than characteristics controllable by the individual. The cladistics analysis also was able to identify some emerging characteristics that

**Table 1.** Relationship between unsecured debt and multiple demographic and personal attributes occurring in combination (OLS regression results).

| Variable | Coefficient | St Err | Beta | t stat | $P$ value | Signif sign |
|---|---|---|---|---|---|---|
| Age | 549.54 | 15.08 | 0.2368 | 36.4480 | 0.0000 | + |
| Female | –3,549.67 | 402.40 | –0.0572 | –8.8210 | 0.0000 | – |
| Couple | 4,264.54 | 700.91 | 0.0678 | 6.0840 | 0.0000 | + |
| Children | –489.60 | 449.26 | –0.0077 | –1.0900 | 0.2758 | |
| Previously bankrupt | –7,295.69 | 506.11 | –0.0896 | –14.4150 | 0.0000 | – |
| Remote rural | 355.84 | 1,889.42 | 0.0012 | 0.1880 | 0.8506 | |
| Metropolitan | 5,493.67 | 422.63 | 0.0818 | 12.9990 | 0.0000 | + |
| Single income | 226.33 | 708.85 | 0.0034 | 0.3190 | 0.7495 | |
| Manager/professional | 11,150.80 | 566.83 | 0.1329 | 19.6720 | 0.0000 | + |
| Clerical/machinery | 2,083.44 | 504.26 | 0.0274 | 4.1320 | 0.0000 | + |
| Labourer | –3,069.74 | 583.59 | –0.0347 | –5.2600 | 0.0000 | – |
| Income | 0.19 | 0.0082 | 0.1540 | 23.1450 | 0.0000 | + |
| (Constant) | –15,477.89 | 1,094.11 | | –14.1470 | 0.0000 | |

Note 1: OLS regression diagnostics: Multiple R 0.3805.
Note 2: Adjusted $R^2$ 0.1448; F-stat 317.55; signif $p = 0.0000$; d.f. 12 and 22,504 (residual).

existed at a lower level than the systemic characteristics, namely, motor vehicle ownership, superannuation and insurance, credit card liabilities, primary income source and gender that may emerge as a systemic characteristic, that is, further socio-economic characteristics related to asset ownership and liabilities may emerge as systemic characteristics. Further analysis considering the macroeconomic factors of the change in GDP, change in interest rates and the change in unemployment rates indicated that the change in interest rates and the change in the unemployment rate were very significant drivers of individual bankruptcies and rank with the micro-economic factors of age and gross income. The use of cladistics then was able to identify that both micro-economic and macroeconomic factors were systemic factors in influencing bankruptcy, and that there were emerging characteristics that needed to be watched, which is not possible with regression type analysis. The cladistics analysis identifies the factors/characteristics that are the most common in affecting bankruptcies, which is a very different concept to regression analysis which is focused on finding weights for various factors such that in aggregate, the outcome of the equation is as close as possible to the observed values and is an "on average" estimate.

## 5.2. Operational risk events

Another interesting example of where cladistics analysis can yield richer insights into financial events is analysis of operational risk events. Chernobai *et al.* (2011) used regression analysis to identify that most operational losses in US banks over the period 1980–2005 could be traced to a breakdown of internal control, and that the banks with greater operational risk events tended to be younger and more complex, have higher credit risk, more anti-takeover provisions and chief executive officers (CEOs) with higher stock option holdings and bonuses relative to salary than other banks. This analysis highlights one of the issues involved with statistical analysis in that it requires factors that can be quantified and that in itself brings issues of relevancy and reliability of the outcomes. A comparable cladistics analysis using the Carmin-Sokal algorithm of US bank operational risk events over the period 2008–2014 by Li *et al.* (2017*a*) identified the

systemic drivers were regulatory issues, multiple people involved, poor controls, legal issue, internal fraud, crime, external fraud, misleading information, computer hacking and "big banks involved" which is significantly different to Chernobai *et al.* (2011). The different results in this instance are primarily due to the ability of cladistics analysis to work with characteristics rather than just statistical data. The Li *et al.* (2017*a*) results are of value to both banks and regulators in terms of trying to manage operational risk events as the cladistics analysis over time was able to also show stability of the systemic drivers, giving management and regulators greater confidence to concentrate on these characteristics alone to reduce operational risk events. Appendix A includes a more detailed explanation of how the cladistics analysis was applied to US operational risk events and is extracted from Li *et al.* (2017*a*).

## 5.3. Motor vehicle insurance claims

Whilst Boucher & Denuit (2006) were primarily concerned with establishing the effects of alternative methodologies for Poisson regression when analysing motor vehicle claims, their analysis showed that categorising young drivers as a homogenous group from a risk perspective was not appropriate and that individual characteristics should be considered. When applied to financial events, and subject to aligning the choice of characteristics to both the research outcome and the algorithm being applied, cladistics analysis will show systemic characteristics and could be applied to analyse the issue being raised by Boucher & Denuit (2006). Whilst not specifically addressing the issue raised by Boucher & Denuit (2006); Evans & Li (2018*b*) used cladistics analysis with the Fitch algorithm and encoding to transform the continuous characteristics into a binary format to determine the systemic characteristics of over 200,000 motor vehicle claims and concluded that whereas the policies had been underwritten with 16 rating factors, there were only 3 systemic characteristics, and importantly, there were another 3 emerging systemic characteristics. Cladistics analysis could be easily used to research further the systemic characteristics of the younger drivers identified in Boucher & Denuit (2006) subject to data availability.

## 6. Conclusion

Given the financial systems are CAS which exhibit emergence through interactions of agents in the financial systems, and move through transition phases, then traditional statistical analysis will only be reliable where the predictive time horizon is very short, and over longer periods should be expected to have a high failure rate due to emergence of the financial system. Cladistics analysis offers an additional basis of analysis for financial systems and will show the systemic drivers of events and the emergence of new systemic characteristic combinations or stability if a temporal analysis is undertaken. For a richer analysis of financial systems, a cladistics analysis should be undertaken to compliment traditional statistical analysis.

## References

Ali, P., Anderson, M., O'Brien, L. & Ramsay, I. (2016). The incidence and causes of personal bankruptcy in Australia. *Finsia Journal of Applied Finance*, (4), 27–35.

Allan, N., Cantle, N., Godfrey, P. & Yin, Y. (2012). A review of the use of complex systems applied to risk appetite and emerging risks in ERM practice. *British Actuarial Journal*, **18**(1), 1–72. https://doi.org/10.1017/S135732171200030X

Allan, N., Yin, Y. & Cantle, N. (2010). Risk DNA: an evolutionary approach to identifying emerging and adapting enterprise risk using phylogenetic analysis. Paper presented at the 2010 Enterprise Risk Management Symposium, Chicago. https://www.soa.org/library/monographs/other-monographs/2010/april/mono-2010-m-as10-1-allan.pdf [accessed 10-Jul-2018].

Archie, J.W. (1985). Methods for coding variable morphological features for numerical taxonomic analysis. *Systematic Biology*, **34**(3), 326–345. https://doi.org/10.1093/sysbio/34.3.326

Arthur, W.B., Durlauf, S. & Lane, D.A. (1997). Introduction: process and emergence in the economy. In W.B. Arthur, S. Durlauf & D.A. Lane (Eds.), *The Economy as an Evolving Complex System II* (Vol. XXVII, pp. 1–14). Addison-Wesley, MA.

Baum, D.A. & Smith, S.D. (2013). *Tree Thinking: An Introduction to Phylogenetic Biology*. Roberts and Company Publishers, Greenwood Village.

Boucher, J. & Denuit, M. (2006). Fixed versus random effects in Poisson regression models for claim counts: a case study with motor insurance. *Astin Bulletin*, **36**(1), 285–301. https://doi.org/10.2143/AST.36.1.2014153

Brown, J.W., Parins-Fukuchi, C., Stull, G.W., Vargas, O.M. & Smith, S.A. (2017). Bayesian and likelihood phylogenetic reconstructions of morphological traits are not discordant when taking uncertainty into consideration: a comment on Puttick et al. *Proceedings of the Royal Society B: Biological Sciences*, **284**(1864), 20170986. https://doi.org/10.1098/rspb.2017.0986

Camin, J.H. & Sokal, R.R. (1965). A method for deducing branching sequences in phylogeny. *Evolution*, **19**(2), 311–326. https://doi.org/10.2307/2406441

Carpenter, J.M. (1992). Random cladistics. *Cladistics*, **8**(2), 147–153. https://doi.org/10.1111/j.1096-0031.1992.tb00059.x

Carpenter, J.M., Goloboff, P.A. & Farris, J.S. (1998). PTP is meaningless, T-PTP is contradictory: a reply to Trueman. *Cladistics*, **14**(1), 105–116. https://doi.org/10.1111/j.1096-0031.1998.tb00206.x

Chernobai, A., Jarion, P., & Yu, F. (2011). The determinants of operational risk in U.S. financial institutions. *Journal of Financial and Quantitative Analysis*, **46**(6), 1683–1725. https://doi.org/10.1017/S0022109011000500

Cilliers, P. (1998). *Complexity and postmodernism : understanding complex systems*. Routledge, London, New York.

Colless, D.H. (1980). Congruence between morphometric and allozyme data for Menidia species: a reappraisal. *Systematic Zoology*, **29**(3), 288–299. https://doi.org/10.2307/2412663

Daníelsson, J. (2008). Blame the models. *Journal of Financial Stability*, **4**(4), 321–328. https://doi.org/10.1016/j.jfs.2008.09.003

Danielsson, J., James, K.R., Valenzuela, M. & Zer, I. (2016). Model risk of risk models. *Journal of Financial Stability*, **23**, 79–91. https://doi.org/10.1016/j.jfs.2016.02.002

Dollo, L. (1893). Les lois de l'évolution. *Bulletin de la Société Belge de Géologie*, **VII**, 164–166.

Evans, J., Allan, N. & Cantle, N. (2017). A new insight into the world economic forum global risks. *Economic Papers: A journal of applied economics and policy*, **36**(2), 185–197. https://doi.org/10.1111/1759-3441.12172

Evans, J. & Li, Y. (2018*a*). A review of global banking regulation under an assumption of complexity. *Annals of Actuarial Science*. http://doi.org/10.1017/S1748499518000301

Evans, J. & Li, Y. (2018*b*). A systems analysis of motor vehicle claims. https://doi.org/10.2139/ssrn.3252224

Faith, D.P. & Cranston, P.S. (1991). Could a cladogram this short have arisen by chance alone?: on permutation tests for cladistic structure. *Cladistics*, **7**(1), 1–28. https://doi.org/10.1111/j.1096-0031.1991.tb00020.x

Farris, J.S. (1973). A probability model for inferring evolutionary trees. *Systematic Zoology*, **22**(3), 250–256.

Farris, J.S. (1983). The logical basis of phylogenetic analysis. In N.I. Platnick & V.A. Funk (Eds.), *Advances in Cladistics. Volume 2, Proceedings of the Second Meeting of the Willi Hennig Society* (pp. 7–36). Columbia University Press, New York.

Farris, J.S., Kluge, A.G. & Carpenter, J.M. (2001). Popper and likelihood versus "Popper*". *Systematic Biology*, **50**(3), 438–444.

Felsenstein, J. (1973). Maximum likelihood and minimum-steps methods for estimating evolutionary trees from data on discrete characters. *Systematic Zoology*, **22**(3), 240–249.

Felsenstein, J. (1978). Cases in which parsimony or compatibility methods will be positively misleading. *Systematic Zoology*, **27**(4), 401–410. https://doi.org/10.2307/2412923

Felsenstein, J. (1981). Evolutionary trees from DNA sequences: A maximum likelihood approach. *Journal of Molecular Evolution*, **17**(6), 368–376. https://doi.org/10.1007/BF01734359

Felsenstein, J. (1988). Phylogenies from molecular sequences: inference and reliability. *Annual Review of Genetics*, **22**, 521–565. https://doi.org/10.1146/annurev.ge.22.120188.002513

Fitch, W.M. (1971). Toward defining the course of evolution: minimum change for a specific tree topology. *Systematic Biology*, **20**(4), 406–416.

Ganegoda, A. & Evans, J. (2014). A framework to manage the measurable, immeasurable and the unidentifiable financial risk. *Australian Journal of Management*, **39**(1), 5–34. https://doi.org/10.1177/0312896212461033

Goldman, N. (1990). Maximum likelihood inference of phylogenetic trees, with special reference to a Poisson process model of DNA substitution and to parsimony analyses. *Systematic Biology*, **39**(4), 345–361. https://doi.org/10.2307/2992355

Goloboff, P.A., Galvis, A.T. & Arias, J.S. (2018). Parsimony and model-based phylogenetic methods for morphological data: comments on O'Reilly et al. *Palaeontology*, **61**(4), 625–630.

Goloboff, P.A., Torres, A. & Arias, J.S. (2017). Weighted parsimony outperforms other methods of phylogenetic inference under models appropriate for morphology. *Cladistics*, **34**, 407–437. https://doi.org/10.1111/cla.12205

Griffiths, A.J.F., Wessler, S.R., Lewontin, R.C., Gelbart, W.M., Suzuki, D.T. & Miller, J.H. (2005). Mutation, repair, and recombination. In *An Introduction to Genetic Analysis* (8th ed., pp. 451–480). W.H. Freeman and Company, New York.

Helfenbein, K.G. & DeSalle, R. (2005). Falsifications and corroborations: Karl Popper's influence on systematics. *Molecular biology and evolution*, **35**(1), 271–280.

Hillis, D.M. (1995). Approaches for assessing phylogenetic accuracy. *Systematic Biology*, **44**(1), 3–16.

Holland, J.H. (1995). *Hidden Order: How Adaptation Builds Complexity*. Addison-Wesley, New York.

Holland, J.H. (2014). *Complexity: A Very Short Introduction*. Oxford University Press, New York.

Kidd, K.K. & Sgaramella-Zonta, L.A. (1971). Phylogenetic analysis: concepts and methods. *The American Journal of Human Genetics*, **23**(3), 235–252.

Kitching, I.J., Forey, P.L., Humphries, C. & Williams, D.M. (1998). *Cladistics: The Theory and Practice of Parsimony Analysis* (Vol. 11). Oxford University Press, New York.

Kluge, A.G. (2006). What is the rationale for 'Ockham's razor' (a.k.a. parsimony) in phylogenetic inference? In V.A. Albert (Ed.), *Parsimony, Phylogeny, and Genomics* (pp. 15–42). Oxford University Press, Oxford, New York.

Kluge, A.G. & Farris, J.S. (1969). Quantitative phyletics and the evolution of Anurans. *Systematic Zoology*, **18**(1), 1–32.

Lake, J.A. (1987). A rate-independent technique for analysis of nucleic acid sequences: evolutionary parsimony. *Molecular biology and evolution*, **4**(2), 167–191.

Lande, R. (1976). Natural selection and random genetic drift in phenotypic evolution. *Evolution*, **30**(2), 314–334. https://doi.org/10.1111/j.1558-5646.1976.tb00911.x

Li, Y., Allan, N. & Evans, J. (2017*a*). An analysis of operational risk events in US and European Banks 2008–2014. *Annals of Actuarial Science*, **11**(2), 315–342. https://doi.org/10.1017/S1748499517000021

Li, Y., Allan, N. & Evans, J. (2017*b*). A nonlinear analysis of operational risk events in Australian banks. *Journal of Operational Risk*, **12**(1), 1–22. https://doi.org/10.21314/JOP.2017.185

Li, Y., Shi, L., Allan, N. & Evans, J. (2018). An analysis of power law distributions and tipping points during the global financial crisis. *Annals of Actuarial Science*, 1–12. https://doi.org/10.1017/S1748499518000088

Matthews, L.J., Edmonds, J., Wildman, W.J. & Nunn, C.L. (2013). Cultural inheritance or cultural diffusion of religious violence? A quantitative case study of the Radical Reformation. *Religion, Brain & Behavior*, **3**(1), 3–15. https://doi.org/10.1080/2153599X.2012.707388

McCarthy, I.P., Tsinopoulos, C., Allen, P. & Rose-Anderssen, C. (2006). New product development as a complex adaptive system of decisions. *The Journal of Product Innovation Management*, **23**(5), 437–456. https://doi.org/10.1111/j.1540-5885.2006.00215.x

Mickevich, M. & Johnson, M.S. (1976). Congruence between morphological and allozyme data in evolutionary inference and character evolution. *Systematic Zoology*, **25**(3), 260–270.

Mitchell, M. (2009). *Complexity: A Guided Tour*. Oxford University Press, New York.

Mitleton-Kelly, E., (Ed.) (2003). *Complex Systems and Evolutionary Perspective on Organisations: The Application of Complexity Theory to Organisations*. Elsevier Science, Oxford.

O'Reilly, J.E., Puttick, M.N., Parry, L., Tanner, A.R., Tarver, J.E., Fleming, J. & Donoghue, P.C.J. (2016). Bayesian methods outperform parsimony but at the expense of precision in the estimation of phylogeny from discrete morphological data. *Biology Letters*, **12**(4), 20160081. https://doi.org/10.1098/rsbl.2016.0081

O'Reilly, J.E., Puttick, M.N., Pisani, D. & Donoghue, P.C.J. (2017). Probabilistic methods surpass parsimony when assessing clade support in phylogenetic analyses of discrete morphological data. *Palaeontology*, **61**(1), 105–118. https://doi.org/10.1111/pala.12330

Pleijel, F. (1995). On character coding for phylogeny reconstruction. *Cladistics*, **11**(3), 309–315. https://doi.org/10.1016/0748-3007(95)90018-7

Popper, K. (1959). *The Logic of Scientific Discovery*. Haper and Row, New York.

Popper, K. (2002). *The Logic of Scientific Discovery*. Routledge, London and Vew York.

Popper, K.R. (1983). *Realism and the Aim of Science*. Hutchinson, London.

Puttick, M.N., O'Reilly, J.E., Tanner, A.R., Fleming, J.F., Clark, J., Holloway, L. & Donoghue, P.C.J. (2017). Uncertain-tree: discriminating among competing approaches to the phylogenetic analysis of phenotype data. *Proceedings of the Royal Society B: Biological Sciences,* **284**(1846). https://doi.org/10.1098/rspb.2016.2290

Queiroz, K.D. (1988). Systematics and the Darwinian Revolution. *Philosophy of Science*, **55**(2), 238–259.

Queiroz, K.D. (2004). The measurement of test severity, significance tests for resolution, and a unified philosophy of phylogenetic inference. *Zoologica Scripta*, **33**(5), 463–473. https://doi.org/10.1111/j.0300-3256.2004.00160.x

Queiroz, K.D. (2014). Popperian corroboration and phylogenetics. *Systematic Biology*, **63**(6), 1018–1022. https://doi.org/10.1093/sysbio/syu064

Queiroz, K.D. & Poe, S. (2001). Philosophy and phylogenetic inference: a comparison of likelihood and parsimony methods in the context of Karl Popper's writing on corroboration. *Systematic Biology*, **50**(3), 305–321.

Rannala, B. & Yang, Z. (1996). Probability distribution of molecular evolutionary trees: A new method of phylogenetic inference. *Journal of Molecular Evolution*, **43**(3), 304–311. https://doi.org/10.1007/BF02338839

Rastogi, S.C., Mendiratta, N. & Rastogi, P. (2008). *Bioinformatics Methods and Applications: Genomics, Proteomics and Drug Discovery*. PHI Learning Private Limited, New Delhi.

Rzhetsky, A. & Nei, M. (1993). Theoretical foundation of the minimum-evolution method of phylogenetic inference. *Molecular Biology and Evolution*, **10**(5), 1073–1095. https://doi.org/10.1093/oxfordjournals.molbev.a040056

Saitou, N. & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, **4**(4), 406–425.

Schrago, C.G., Aguiar, B.O. & Mello, B. (2018). Comparative evaluation of maximum parsimony and Bayesian phylogenetic reconstruction using empirical morphological data. *Journal of Evolutionary Biology*, **31**(10), 1477–1484. https://doi.org/10.1111/jeb.13344

Schweitzer, F., Fagiolo, G., Sornette, D., Vega-Redondo, F., Vespignani, A. & White, D.R. (2009). Economic networks: the new challenges. *Science*, **325**, 422–425.

Shi, L., Allan, N., Evans, J. & Yun, Y. (2018*a*). Significance of controllable and uncontrollable drivers in credit defaults. *Economic Papers: A Journal of Applied Economics and Policy*, **37**(1), 30–41. https://doi.org/10.1111/1759-3441.12200

Shi, L., Evans, J. & Li, Y. (2018*b*). A systems analysis of drivers of individual bankruptcies. *Economic Papers: A Journal of Applied Economics and Policy*, published online August 2018. https://doi.org/10.1111/1759-3441.12229

Siddall, M.E. & Kluge, A.G. (1997). Probabilism and phylogenetic inference. *Cladistics*, **13**(4), 313–336. https://doi.org/10.1111/j.1096-0031.1997.tb00322.x

Smith, A. (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations*. The Electric Book Co, London.

Sober, E. (1975). *Simplicity*. Clarendon Press, Oxford.

Sokal, R.R. & Michener, C.D. (1958). A statistical method for evaluating systematic relationships. *The University of Kansas Science Bulletin*, **38**, 1409–1438.

Song, F. & Thakor, A.V. (2010). Financial system architecture and the co-evolution of banks and capital markets. *The Economic Journal*, **120**(547), 1021–1055.

Studier, J.A. & Keppler, K.J. (1988). A note on the neighbor-joining algorithm of Saitou and Nei. *Molecular Biology and Evolution*, **5**(6), 729–731. https://doi.org/10.1093/oxfordjournals.molbev.a040527

Thiele, K. (1993). The Holy Grail of the perfect character: the cladistic treatment of morphometric data. *Cladistics*, **9**(3), 275–304. https://doi.org/10.1111/j.1096-0031.1993.tb00226.x

ul-Haq, R. (2005). *Alliances and Co-Evolution: Insights from the Banking Sector*. Palgrave Macmillan, New York.

Wiens, J.J. (2001). Character analysis in morphological phylogenetics: problems and solutions. *Systematic Biology*, **50**(5), 689–699.

Wiens, J.J., (Ed.) (2000). *Phylogenetic Analysis of Morphological Data*. Smithsonian Institution Press, Washington and London.

Wiley, E.O. & Lieberman, B.S. (2011). *Phylogenetics: Theory and Practice of Phylogenetic Systematics*. John Wiley & Sons, New Jersey.

Willson, S.J. (1999). A higher order parsimony method to reduce long-branch attraction. *Molecular Biology and Evolution*, **16**(5), 694–705. https://doi.org/10.1093/oxfordjournals.molbev.a026152

Wright, A.M. & Hillis, D.M. (2014). Bayesian analysis using a simple likelihood model outperforms parsimony for estimation of phylogeny from discrete morphological sata. *PLoS ONE*, **9**(10), e109210. https://doi.org/10.1371/journal.pone.0109210

Yang, Z. & Rannala, B. (1997). Bayesian phylogenetic inference using DNA sequences: a Markov Chain Monte Carlo Method. *Molecular Biology and Evolution*, **14**(7), 717–724. https://doi.org/10.1093/oxfordjournals.molbev.a025811

## Appendix A

**An application of cladistics analysis to US bank operational risks[2]**

We extracted data from the ORIC International[3] database of operational risk events for US banks from 2008 to the middle of 2014 and after filtering and cleaning the data, the US data contained 1371 unique risk events. We derived from the data a set of characteristics as shown in Appendix B[4] to use in the analysis. Several software programs are able to perform cladistics analysis using the maximum parsimony algorithm, and in this study, we used software from Systemic Consult[5]. Alternative software is available in R. Figure A1 presents an example of the output of the cladistics analysis. This tree is read from left to right. The left most characteristic, that is, "Internal fraud", can be thought of as the "level 1" characteristic and occurs for a group of risk events. These level 1 characteristics are important, as whilst all the characteristics leading to a risk event must be linked for that risk event to occur, if an institution can break the chain of linking characteristics, then the risk event won't occur. Given the level 1 characteristics are those characteristics that are common to several risk events, then it is logical to concentrate on managing these systemic characteristics to mitigate risk events occurring. The level 2 characteristics, that is, "Multiple people" and "Single person" are not as systemic as the level 1 characteristic, and the level 3 characteristics in this figure, that is, "Poor controls", "Crime" and "Big banks involved" are the un-systemic characteristic for each event. One characteristic can appear in different places and at different levels, that is, in this tree, "Crime" appears at level 3 for different events. The value of applying cladistics analysis to financial events is the ability to derive the level 1 characteristics and establish their stability, as that is a necessary prerequisite to their being an efficient target for risk mitigation. Tables A1 and A2 show the level 1 characteristics we derived for the US banks for both cumulative and independent periods.

---

[2] Extracted from Li *et al.* (2017a)

[3] https://www.oricinternational.com/

[4] The derivation of the characteristic set is subjective, but involves using word counting software to determine the most common descriptions of events as reported which then need to be adjusted to meet the criteria for cladistics analysis that the selected characteristics either were involved or were not involved for each event.
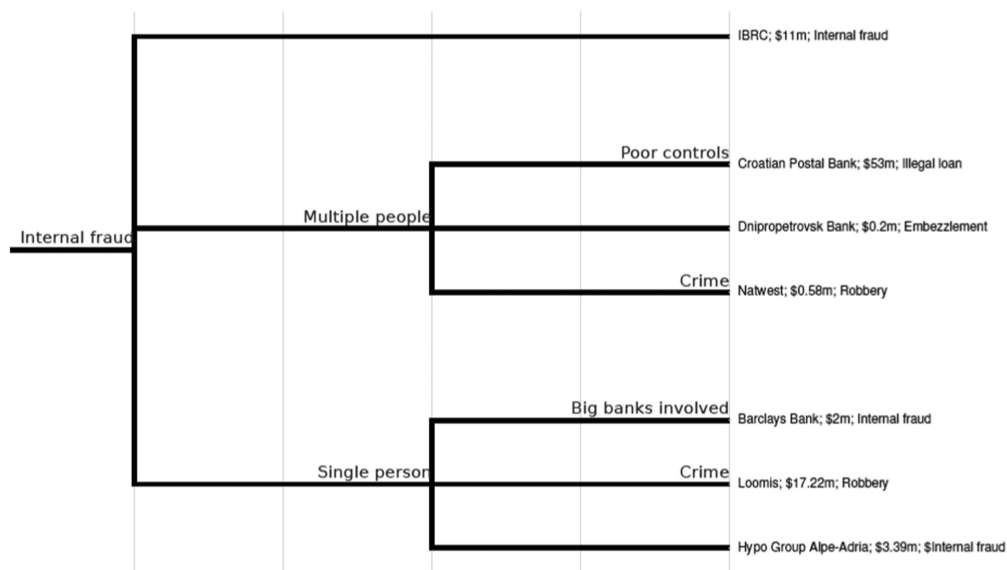
[5] http://www.systemicconsult.com/

374

**Figure A1.** An example of cladistics tree for US Bank Operational Risks.

**Table A1.** Level 1 characteristics for cumulative periods for US banks.

|  | 2008–2010 | 2008–2011 | 2008–2012 | 2008–2013 | 2008–2014 |
|---|---|---|---|---|---|
| Regulatory issues |  | x | x | x | x |
| Multiple people | x |  | x |  |  |
| Poor controls | x | x | x |  |  |
| Legal issue |  | x | x | x | x |
| Internal fraud |  | x | x | x | x |
| Crime |  |  |  |  |  |
| External fraud |  | x | x | x | x |
| Single person | x |  |  |  |  |
| Misleading information | x | x |  |  |  |
| Derivatives |  |  |  |  |  |
| ATM |  |  |  |  |  |
| Complex transaction |  |  | x |  |  |
| Computer hacking | x |  |  | x | x |
| Complex products |  |  |  |  |  |
| Money laundering |  |  |  |  |  |
| Software issue |  |  |  |  |  |
| International transaction |  |  |  |  |  |
| Overcharging |  |  |  |  |  |
| Credit card |  |  |  |  |  |
| Employment issues |  |  |  |  |  |
| Insurance |  |  |  |  |  |
| Manual process |  |  |  |  |  |
| Bank cross-selling |  |  |  |  |  |
| Offshore fund |  |  |  |  |  |
| Big banks involved | x | x | x | x | x |

Note 1: The significant level 1 characteristics are "regulatory issues," "legal issues," "internal fraud," "external fraud" and "big banks involved."

**Table A2.** Level 1 characteristics for independent periods for US banks.

| | 2008–2010 | 2011–2012 | 2013–2014 |
|---|---|---|---|
| Regulatory issues | | x | x |
| Multiple people | x | | x |
| Poor controls | x | x | x |
| Legal issue | | x | x |
| Internal fraud | | x | x |
| Crime | | x | x |
| External fraud | | x | x |
| Single person | x | | |
| Misleading information | x | | x |
| Derivatives | | | |
| ATM | | | |
| Complex transaction | | | |
| Computer hacking | x | x | |
| Complex products | | | |
| Money laundering | | | |
| Software issue | | | |
| International transaction | | | |
| Overcharging | | | |
| Credit card | | | |
| Employment issues | | | |
| Insurance | | | |
| Manual process | | | |
| Bank cross-selling | | | |
| Offshore fund | | | |
| Big banks involved | x | x | x |

Note 1: The significant level 1 characteristics are "regulatory issues," "multiple people," "poor controls," "legal issue," "internal fraud," "crime," "external fraud," "misleading information," "computer hacking" and "big banks involved." From Tables A1 and A2 we can observe that the level 1 characteristics that emerge for the US market are "big banks involved," "poor controls," "regulatory issues," "legal issues," "internal fraud" and "external fraud." These characteristics not only emerge in each independent period but also for the entire period and suggest US banks, in both their daily management and business activities, are weak in process control, and historically, may not have paid sufficient attention to regulations or had weak compliance processes. The value this analysis brings is that it identifies sustainable systemic drivers of operational risks, providing banks with clear direction as to where they need to improve their operational processes if they wish to reduce operational losses.

# Appendix B: Characteristics used in the US bank operational risk analysis

**Table B1.** Characteristics used in the US bank operational risk analysis

|   | Characteristic | Definition |
|---|---|---|
| 1 | Poor controls | Event where controls that should have been in place were not or were ineffective |
| 2 | Single person | Event initiated by an individual |
| 3 | Crime | Event involving theft other than by deception |
| 4 | Internal fraud | Event involving fraudulent activity by a member of staff |
| 5 | External fraud | Event involving fraudulent activity by an external person(s) |
| 6 | Multiple people | Event imitated by many people |
| 7 | Regulatory failure | Event where a government regulation was breached |
| 8 | International transaction | Event involving a transaction occurring across a country border |
| 9 | ATM | Event involving an ATM |
| 10 | Complex transaction | Event involving a transaction that involved many parts |
| 11 | Legal issue | Event where a customer took an institution to court for remedy, but the event was not a regulatory breach |
| 12 | Credit card | Event involving use/misuse of a credit card |
| 13 | Human error | Event where a staff member made a mistake |
| 14 | Misleading Information | Event where the product/service details were not made clear to a customer |
| 15 | Complex products | Event involving products that had numerous components |
| 16 | Bank cross-selling | Event involving a bank selling a product/service to a customer that was different to what the customer originally bought from the bank |
| 17 | Overcharging | |
| 18 | Employment issues | Event where employment contract conditions or government regulations relating to employment were breached |
| 19 | Computer hacking | Event involving hacking into a system |
| 20 | Manual process | Event involving a manual process |
| 21 | Offshore fund | Event where a transaction involved a fund that was domiciled outside the country where the investor was located |
| 22 | Money laundering | Event where funds were transferred for the purposes of creating a false impression that the transaction was legitimate |
| 23 | Software system | Event involving a software issue |
| 24 | Insurance | Event involving an insurance product |
| 25 | Derivatives | Event involving a derivative transaction |
| 26 | Big banks involved | Event involving big banks |