A Social Theory of Networks in Finance: Financial Careers, Revolving Doors and the Great Financial Crisis

Cullen, Conor Fullerton, Max Scibetta, Alessandro Zulkarnaen, Farrel

May 12, 2019

Project Supervisor
Dr. Daniel Tischer

Abstract

The events of the 2008 Great Financial Crisis (GFC) contributed to the greatest shock to the financial system in almost a century. In the years following the Crisis, the over-the-counter derivatives traded via the complex shadow banking system were brought to light, and their impact on the demise of the global markets has been widely investigated. One of the most commonly traded securities in the years leading up to the GFC was the Collateralised Debt Obligation (CDO), observed to be the engine which drove the subprime mortgage crisis. Guided by the question of whether there are collateral manager career paths that can be identified which help understand CDO performance, this report investigates the elaborate networks of actors and firms involved during the lifetime of a CDO. The performance of CDOs and the network of associated actors and firms is analysed through implementation of Exponential Random Graph Modelling and the Quadratic Assignment Procedure. These techniques allow for the significance of any observed network characteristics to be quantified. In line with current understanding, we show that collateral managers with particular expertise in investment banking and credit ratings agencies have the most significant impact on CDO performance. We also show that firms comprised of actors with careers of significant tenure at particular firms have a generally negative influence on the performance of their respective CDOs.

Contents

1 Introduction

Following the Great Financial Crisis (GFC) of 2008, the Collateralised Debt Obligation (CDO) was highlighted as one of the exotic financial derivatives involved in the collapse of the financial sector. Whilst there is a large body of work regarding the CDO and its role in the collapse, far less is known of the complex network of financial institutions underpinning these derivatives. Arguably, the most crucial actors in this network were a set of firms involved in the structurisation and management of the CDO, known as Collateral Management Firms (CMFs). As independent, third party firms, the CMFs were employed to select and manage the assets within a CDO, with a fiduciary duty to keep their investor's best interests in mind. However, in a stark conflict of interest, they were hired by the investment banks whose CDOs they managed.

Social network theory presents a method of analysing the previous career movements of people who worked in CMFs. A company can be represented as having connections to employees if they have worked there at some point in their career. The aim of this project is to discover whether the connections have any meaningful relationship to how CDOs were downgraded during the Financial Crisis. To achieve this aim, a series of objectives are laid out as follows:

1. Measuring CDO performance through a detailed score system based on the reported credit ratings.
2. Analysing social sequence patterns to discover if specific career patterns are present and what consequences they have on CDO performance.
3. Developing social networks for the individuals whose careers were directly related to both the CMFs and banks associated with a CDO.
4. Analysing these networks to observe any structural features that may be correlated with the calculated average CDO performance ratings.

2 Background

2.1 What is a CDO?

In simple terms, a CDO is a structured asset backed security. Securitisation is a financial practice that pools various types of debt like loans, bonds and mortgages and sells the related cash flows generated from the interest and principal repayments to third party investors for a lump sum. This allows the transfer of risk to be traded from the originator of the illiquid debt to the investor and frees up the issuer financially to write more loans. The promised payments from these debts are the collateral that gives the CDOs their value, hence the name collateralised debt obligation. A CDO is a structure that guarantees to pay investors in a prescribed way by splitting the financial claims on the underlying collateral into different levels of seniority, called tranches, displayed in Fig. 1. Credit risk is typically assessed based on the probability of default of the assets, with the safest occupying the most senior tranches and the riskiest the most junior. The structured nature of these products allow for risk to be redistributed, making it possible for investors with different risk appetites to participate in the market. The safest tranches have first claim on the cash flows but equally receive the smallest returns. The remaining funds then cascade down with each tranche's payments being fulfilled before passing onto the next. If insufficient funds are available to the CDO then the most junior tranche will suffer the first losses. This effect allows the riskier tranches to act as a protective layer to those above.

2.2 Network Agents

From the point of origination to when an investor ends up with a borrower's monthly repayments as part of a CDO, there are several different institutions that all carve out some percentage of the initial fees or monthly cash flow. This report focuses on the set of firms that were involved in the structurisation of the CDO and the resulting inter-organisational networks that were formed. There are a myriad of financial and legal actors that participated in the creation of a CDO including accountants, attorneys, collateral administrators, credit rating agencies and trustees. However, the two main entities of concern in this report will be that of the underwriter, typically an investment bank, and the CMF. The CMF's job is to help structure the securities and work with the underwriter to sell them to investors. If the transactions of the securities are properly structured and the pools perform as expected then the credit risk of all tranches is improved. However, if the pools are improperly structured then the tranches may experience dramatic credit deterioration and loss. The CMF is supposed to be an independent advisor that performs a key risk mitigating function in the creation of a CDO to assuage investor concerns of moral hazard and has a fiduciary duty to protect the assets within.

2.3 Market Behaviour

Investment banks were seeking a way to connect the vast reservoirs of cash that they managed for institutional and foreign investors to the rising property market. Usually these investors only pursue opportunities of the safest kind, known as Aaa rated, such as long-term treasury bonds issued by the Federal Reserve. However, the introduction of the CDO allowed what were normally riskier products to be packaged up and sold as largely Aaa investments with much higher returns than treasury bonds, sparking huge demand on the secondary mortgage market. Loan originators earned fees not based on quality but quantity and with growing demand for their products they began lowering their underwriting standards and adopting predatory practices. When credit worthy customers had been exhausted, attention then turned to those that had subprime financials. Mortgage lenders approved loans without fully inspecting their customer's ability to pay, allowing many borrowers to take on loans larger than they could afford with the hope that they could refinance or sell their homes at a higher price and profit from the rise in equity later on.

The Federal Reserve, noticing an overheating economy with unsustainable property prices, aimed to correct the market and continuously drove interest rates back up. Demand slowed and with a surplus in supply from new home developments, prices started to plummet. Borrowers were left with rising mortgage payments that they could no longer afford and houses worth less than the mortgage they had taken out to finance the purchase. Default rates began skyrocketing and the value of the underlying assets of the CDOs were left unknown which began the ensuing collapse of the mortgage and debt market. This triggered the banking crisis and then what was known as the GFC.

3 Methods

3.1 Data

The deregulation of banks and the rules surrounding the trading of derivatives allowed the financial institutions creating CDOs to cloud their documents in mystery. They are mundane, highly structured and both technically and legally convoluted products that are not presumed to be read by external parties. There is also no public database that exists containing all documents so no conclusive figures for the size of the CDO market leading up to the financial crises. However, a document called the Offering Circular was issued by banks for each CDO as a means to advertise the deal to investors. A database collating the technical information within these documents, sourced from a wide range of places but primarily the Irish Stock Exchange, was populated and shared for the purpose of this project. The information includes descriptions of the product's structures, management and processes. The information of particular interest in these documents is the career history of those involved with the product from a CMF perspective. The sequence and tenure of the roles for each individual provide a sample for the expertise and networks that contributed to the performance of these products. Research was carried out to determine the type of company an individual gained experience from in order to construct an overview of the differing social sequences and analyse any potential patterns.

In addition to this information is the investment quality ratings issued by the third party Credit Rating Agencies. These provide a temporal axis along which performance can be judged. Assessments of the credit risk occur at discrete moments and the resulting downgrades can be collated to provide a performance over the lifteime of the product.

3.2 Rating CDO Performance

The dataset contains the history of recorded credit ratings (given on the Moody's long-term credit rating scale) and their respective dates for each tranche of 55 individual CDOs. The Moody's scale gives a rating of opinion of risk on a fixed-income obligation, using an alphanumeric score on a scale of size 21 (as shown in the first column of Tab. 1). The highest quality senior tranches, deemed to have the most minimal risk are usually rated Aaa, whilst the poorest quality junior tranches are usually rated C, and are typically in default [1]. There is also the case where a CDO tranche is rated W. This rating is given in the event that the credit rating agency no longer rates the obligation in question. This may happen for a number of reasons, namely bankruptcy of the CDO issuer or misinformation at the time of rating [2]. Many of the CDOs in the dataset have numerous rating change events, on multiple different dates in the time between their initial issuance and the height of the financial crisis in September 2008. Some CDOs even continued to be downgraded until 2010. As a consequence, directly comparing the credit rating performance of each CDO, over the course of multiple downgrade events is not practically possible without the use of a specific metric.

To generate a metric, each tranche is assigned a numerical value based on its rating on the Moody's scale, Aaa being assigned 21, C assigned 1 and W being assigned 0 (see Tab. 1). The initial score of each tranche in a CDO is summed, giving each CDO an initial overall score, which is referred to as its performance rating. Each downgrade event allows the calculation of the difference in performance rating between the most recent rating and that prior to it. For example, a downgrade from Aaa (assigned 21) to A2 (assigned 16) would correspond to a downgrade score of 5. The sum of the downgrade scores gives a value magnitude of default (an unscaled measure of the poorness of performance of a given CDO).

3.3 Sequencing and Cluster Analysis

The dataset includes the career patterns (termed as sequences) of 553 individual 'actors'. A sequence is defined here as a succession of elements chosen inside of an alphabet. In other words, a sequence is a time series with categorical data. In this case, the categories are the types of firm which an individual has worked at over their career. The succession follows the order of time and describes an individual trajectory of a person's career (e.g. investment bank for 15 years, followed by asset management for 10 years, etc.) up until the year the GFC began. Social sequences are made of three basic dimensions: the order in which they occur; the nature of the successive states, chosen among the alphabet; and their duration (i.e. the duration of constant sub-sequences) [3].

Cluster analysis is a useful technique used for social sequence analysis [4]. The algorithm used here includes optimal matching [5]. Let $S = (s_1, s_2, ..., s_T)$ be a sequence of states s_i belonging to a finite set of possible states in the alphabet, where T is the size of the number of sequences, with $1 \leq i \leq T$. It is also known that $S \in \mathcal{S}$, where $\mathcal{S}$ is the sequence space, the set of all possible sequences of states. The maximum optimal matching distance for two states s_i and s_j for $1 \leq i, j \leq T$ is defined as

$$D_{max} = \min(\ell_x, \ell_y) \cdot \min(2c_I, \max(SC)) + c_I|\ell_x - \ell_y|, \tag{1}$$

where ℓ_x and ℓ_y are the sequence lengths, c_I the indel cost (i.e. the cost of inserting an element in the sequence) and SC is the substitution cost matrix [6]. Here the indel cost is set at half the maximum substitution cost, meaning that it will never cost less to make an insertion and a deletion in place of a substitution. The substitution cost matrix has dimensions TxT and the element (i, j) is the cost of substituting state i with state j. Clustering is then performed by means of a hierarchical, ascending (or agglomerative) model with square Euclidean distances and Ward's aggregation algorithm [7]. This uses the method where the criterion for choosing a pair of clusters to merge at each step is based on the optimal value of an objective function. The objective function is defined here as the square Euclidean distance,

$$d_{ij} = d(\{X_i\}, \{X_j\}) = ||X_i - X_j||^2, \tag{2}$$

where X_i and X_j are two adjacent clusters for the elements (i, j) of the matrix SC.

3.4 Network Dynamics

The management careers compose a very large bipartite network with a population size of over 1000 unique actors. A bipartite (or two-mode) network is a set of graph nodes decomposed into two disjoint sets such that no two nodes in the same set are connected. The two node sets are actors and firms. In set notation, this translates to a graph $G = \{U, V, E\}$, where U and V are the set of actor nodes and company nodes respectively, and E is the set of edges. For a node u or v in graph G where either $u \in U$ or $v \in V$, the number of adjacent nodes is called the degree of the vertex and is denoted by $\deg(u)$ or $\deg(v)$. The degree sum formula of a bipartite graph states that

$$\sum_{v \in V} \deg(v) = \sum_{u \in U} \deg(u) = |E|, \tag{3}$$

where $|E|$ denotes the cardinality of the set E. In order to analyse the graph in a meaningful way, the information from the bipartite network is compressed in the form of a network projection on the set of nodes U (call this U'). In this projected network, only nodes in the set U (actors) are included in the graph, and a connection between two nodes corresponds to both nodes having at least one common node in the set V (termed as being adjacent). This translates to two company nodes being connected if an actor has worked at both companies.

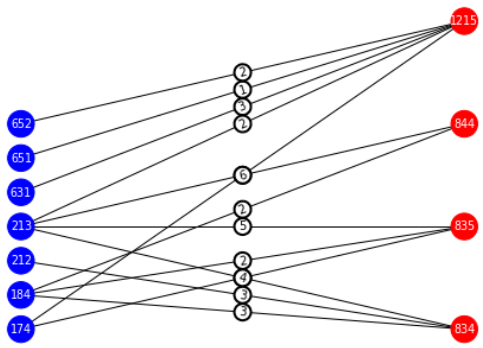

Figure 2a: Example of a bipartite graph. In particular, this is the ego network (see Section 3.8) of the credit rating agency Fitch. Blue nodes are people IDs (node set U) and red nodes are company IDs (node set V). Edge weights correspond to the tenure at a company.

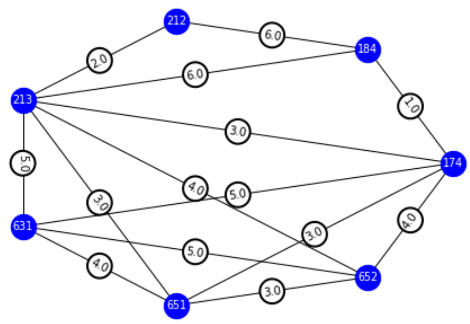

Figure 2b: Projection on the node set V in the bipartite graph from Fig. 2a. Weights are calculated according to Eq. 4.

As the original bipartite graph has weighted edges, the redistribution of weights must be carefully chosen as it is said to have a strong effect on the community structure of the overall network [8]. As the weighting system is based on the number of years worked at a company (tenure), the best way to redistribute the weights is to take the average weighted degree as the new weight. For two adjacent nodes $u_i \in U$ and $u_j \in U$ with n edges connected to nodes $v_k \in V$ that have unique edge weights w_k, the new weight w'_{ij} in the projected network is

$$w'_{ij} = \frac{\sum_k^n w_k}{n}, \tag{4}$$

where $i, j \in [0, |U|]$ and $k \in [0, n]$. The adjacency matrix A of the graph U' is symmetric and defined such that its element A_{ij} is w'_{ij} when there is an edge from node u_i to u_j, and 0 when there is no edge,

$$A_{ij} = \begin{cases} w'_{ij} & \text{if connected,} \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

To analyse this graph, one must filter the network to include only firms of interest which are responsible for the management of certain CDOs. The way to do this is by using ego networks of these firms. The ego network consists of a focal node ('ego') and nodes that are directly connected to it ('alters') plus the ties, if any, among alters. There is also a neighbourhood of depth N included in the ego network, which is the collection of the ego and all nodes to whom the ego has a connection at a path length of N. In other words, this includes the neighbours of the alters if the path length is two.

3.5 Weighted Degree Distribution

In an undirected network, the degree distribution is defined as the probability distribution of the degrees of each node over the whole network. Most social networks have degree distributions that are highly skewed to the right, meaning a large number of nodes have low degree with a few high degree nodes. From the adjacency matrix defined in Eq. 5, the weighted degree of a node is simply the sum of the weighted edges. In terms of the adjacency matrix A from Eq. 5, the weighted degree of node i, denoted k_i, is the sum of the ith row of A,

$$k_i = \sum_j a_{ij}, \tag{6}$$

where the sum is over all nodes in the network. The weighted degree distribution, $P_{deg}(k)$, is defined as the fraction of nodes in the network with weighted degree k. Since the weights are in terms of average tenure at a firm, the weighted degree distribution can provide information on how careers are distributed across the firms in the bipartite network.

3.6 Exponential Random Graph Modelling

Exponential Random Graph Modelling (ERGM) presents a way to understand the processes of (social) network structure emergence and tie formation. Tie formation is characterised by an 'actor' working for a number of firms specialising in investment banking, credit ratings, etc. The model works by measuring a set of known statistics from a network and uses the exponential family distributions of the statistics to generate random networks. These are then compared to the observed networks to assess the likelihood of fit. This process is used because the observed network is only one instance of a large number of alternative networks which may display similar or dissimilar structural features. For a randomly generated graph Y, the probability of observing a graph y on a fixed set of nodes [9] is

$$P(Y = y|\boldsymbol{\theta}) = \frac{\exp(\boldsymbol{\theta^T g(y)})}{k(\boldsymbol{\theta})}, \tag{7}$$

where $\boldsymbol{g(y)}$ is the vector of arbitrary network statistics, $\boldsymbol{\theta}$ is the vector of model parameters and $k(\boldsymbol{\theta})$ is the numerator summed over all possible networks on node set y. This is also denoted as

$$k(\boldsymbol{\theta}) = \sum_{y \in \mathcal{Y}} \exp(\boldsymbol{\theta^T g(y)}), \tag{8}$$

where $\mathcal{Y}$ is the sample space of Y (for example, $\{0, 1\}^N$ for an unweighted network) [9]. For an undirected network of size n, the cardinality of the set of all possible networks is $2^{n(n-1)/2}$, which is the power set of the number of dyads (potential ties). This means there is a significant blow up of the number of possible networks as n increases; thus, the ERGM works best for relatively small networks (hence the decision is made to consider only the best 5 and worst 5 performing CDOs in Section 4.3). The technique used for generating samples from the set of possible networks is a form of Monte Carlo Markov Chain (MCMC) sampling. This refers to the process of exploring the space of possible networks and moving towards ones which are more likely given the chosen distribution coefficients (depending on the vector $\boldsymbol{\theta}$). The way to generate through random samples is to use the Gibbs sampler [10], cycling through the set of random variables $Y_{ij} (i \neq j)$ and simulating each individually according to the conditional distribution

$$P_\theta\{Y_{ij} = y_{ij}|Y_{uv} = y_{uv}\} \;\; \forall (u, v) \neq (i, j). \tag{9}$$

Continuing this for a large number of samples (at least 10,000) defines a Markov chain on the space of all adjacency matrices converging to the desired stationary distribution. There are other possible sampling methods to approximate the distribution but the one described above is used in the computations to follow.

The type of network statistic that can be used in the model is quite large, and many metrics will not be important to the context of social interactions in the careers network. Defining the network as undirected narrows down the possible features that may be of relevance to CDO performance. The frequency of interaction or density of the graph is a very important aspect of the ERGM, as it relates directly to tie formation. Nodal attributes of the actors are also important in this model, and account for assortativity effects, which is the preference for a network's nodes to attach to others which share a similarity. Homophily is the measure of bias in favour of connections with these similarities. Finally, another measure to consider is the triad closure bias. This is the bias that models connections displaying a 'friend of a friend' effect (see Fig. 3a).

With the network statistics defined, the probability distribution in Eq. 7 becomes [11]

$$P_\theta(Y = y) = \exp\left(\sum_{k=1}^{n-1} \theta_k S_k(y) + \tau T(y) - \psi(\theta, \tau)\right), \; y \in \mathcal{Y}, \tag{10}$$

where the statistics S_k and T are defined by

$$S_1(y) = \sum_{1 \leq i < j \leq n} y_{ij} \;\; \text{number of edges}, \tag{11}$$

$$S_k(y) = \sum_{1 \leq i \leq n} \binom{y_{i+}}{k} \;\; \text{number of } k\text{-stars } (k \geq 2), \tag{12}$$

$$T(y) = \sum_{1 \leq i < j < h \leq n} y_{ij} y_{ih} y_{jh} \;\; \text{number of triangles}, \tag{13}$$

with θ_k and τ as the model parameters and $\psi(\theta, \tau)$ as a normalising constant [11]. The variable y_{i+} denotes the degree of node i. For a structure with central node i and neighbours $j_1, ..., j_k$ where $k \geq 2$, the structure is called a k-star. The quantity $\binom{y_{i+}}{k}$ is the number of k-stars which node i is involved in. Examples of these kinds of structures are shown in Fig. 3.

The conditional log odds of the tie formations is used to estimate the complete vector of statistics. The log odds of a specific tie is defined as

$$\text{logit}(P_\theta\{Y_{ij} = 1 | Y_{uv} = y_{uv}\} \;\; \forall(u,v) \neq (i,j)) = \log\left(\frac{P_\theta\{Y_{ij} = 1 | Y_{uv} = y_{uv}\} \;\; \forall(u,v) \neq (i,j)}{P_\theta\{Y_{ij} = 0 | Y_{uv} = y_{uv}\} \;\; \forall(u,v) \neq (i,j)}\right) = \boldsymbol{\theta^T \partial(g(y))}, \tag{14}$$

where $\boldsymbol{\partial(g(y))}$ is the change in $\boldsymbol{g(y)}$ when Y_{ij} switches between 0 and 1 [12].

The log odds are calculated according to Eq. 14 and estimated by MCMC as it converges towards a maximum likelihood estimate. These estimates can be converted into probabilities ($p(y)$, equivalent to Eq. 10) by the relation between probability and the logit function,

$$p(y) = \frac{\exp(\ln(\text{odds}))}{1 + \exp(\ln(\text{odds}))}, \tag{15}$$

where odds is the estimate calculated in Eq. 14. The plot showing this relation is given in Fig. 4. The probability is equivalent to observing a network y on the fixed set of nodes, with the given set of relationships (i.e. either edges, triangles or k-stars).

3.7 Quadratic Assignment Procedure

One other thing to consider in inference analysis is the confidence in the results. In a general case, inferential statistics can be described as a measure of confidence that the patterns found in the analysis are not by chance but rather are truly sampled from a meaningful population - the apparent pattern is not a random occurrence. However, the problem lies that there is only one sample of data and so there are no other trials to refer to. It is possible, however, to create an empirical sampling distribution through permutation of the data. A common procedure of doing so in the context of social networks is through the use of the Quadratic Assignment Procedure (QAP) [13]. QAP is a non-parametric algorithm for inference in social network analysis specifically for use with dyadic data (pairs of data). In the context of this problem it is a useful way to determine confidence in the similarity between pairs of actors in the networks. QAP returns two statistical values:

- p-value as metric for test if coefficients parameters (βs) are significant.
- ρ correlation coefficient as a measure of similarity between pairs of data.

This is used to infer how similarities between actors can affect CDO performance. As inference, one can view QAP as both a regression and correlation problem.

QAP can be seen as an extension to a multiple regression model (typically coined as Multiple Regression QAP, or MR-QAP). To understand this concept, a reminder of a traditional regression setting is useful. Also, it is important to consider how it may differ for regressions pertaining to dyadic data in networks. In a traditional regression, the most common example of regression problem, Ordinary Least Squares (OLS) is typically modelled as the following:

$$\mathbf{Y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + +\beta_i X_i + \epsilon_i, \tag{16}$$

where $\mathbf{Y}$ is the dependent variable matrix, β represents the coefficients, X_i represents the independent variables for $0 \leq i \leq n$, and ϵ_i represents the residuals, which are assumed to follow a Gaussian distribution. Also, here we denote n is equal to the number of nodes in the network.

However, in MR-QAP regression, each independent variable is now a matrix. Previously, traditional regression describes behaviour of individual observations of cases as single variables, whereas now each cell in the matrix is a single observation. Thus, QAP regression tries to find the patterns or similarities in the relationship between the cells in each matrix (i.e. $\mathbf{X}_{i,j}$), because they are dyadic data the matrix should be symmetric. Now $\mathbf{Y}$ is an $N \times N$ dependent matrix and $\mathbf{X_i}$ are independent matrices. However, because it is dyadic, we can no longer assume that pairs of observations are independent of another. This is the main advantage of QAP. QAP randomises the dependent variables by permutations, as a result this creates random data sets that we can use as an empirical sampling distribution. Through the use of these sampling distributions, it is possible to infer the nature of the coefficients β_i from all the permuted data sets and how statistically significant they are.

QAP acts in two steps. First, it performs standard multiple regression across the cells of the dependent and independent matrices. Second, it randomly permutes the rows and columns of the dependent matrix simultaneously. In order to get a better estimate of the standard errors, this step is repeated several hundred times; hence, the null hypothesis is represented by the permuted dataset. Note, because the data has been randomised, the values pertaining to the rows and columns have been separated. Thus, QAP permutes the rows and columns of the dependent matrix only. With this, the dependent variable values have been separated from the corresponding independent variables.

For example, permuting 1←3, 2←2, 3←4, and 4←1 [13] gives

$$\mathbf{Y} = \begin{pmatrix} y_{1,1} & y_{1,2} & y_{1,3} & y_{1,4} \\ y_{2,1} & y_{2,2} & y_{2,3} & y_{2,4} \\ y_{3,1} & y_{3,2} & y_{3,3} & y_{3,4} \\ y_{4,1} & y_{4,2} & y_{4,4} & y_{4,4} \end{pmatrix} \mapsto \mathbf{Y}_{permuted} = \begin{pmatrix} y_{3,3} & y_{3,3} & y_{3,4} & y_{3,1} \\ y_{2,3} & y_{2,2} & y_{2,4} & y_{2,1} \\ y_{4,3} & y_{4,3} & y_{4,4} & y_{4,1} \\ y_{1,3} & y_{1,3} & y_{1,4} & y_{1,1} \end{pmatrix} \tag{17}$$

As a result, the dependence within rows and columns are preserved, whilst simultaneously removing the relationship between dependent and independent variables.

QAP can also be seen as an auto-correlation problem. That is, the error terms can be assumed to be auto-correlated. Therefore, one way to represent network data is by assuming that they have an auto-correlation matrix which is outputted by QAP [13].

$$\Omega_{ij,kl} = \sigma^2 \begin{pmatrix} 1 & \rho_{12,13} & \rho_{12,14} & \cdots & \rho_{12,n(n-1)} \\ \rho_{13,12} & 1 & \rho_{13,14} & \cdots & \rho_{13,n(n-1)} \\ \rho_{14,12} & \rho_{14,13} & 1 & \cdots & \rho_{14,n(n-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{n(n-1),12} & \rho_{n(n-1),13} & \rho_{n(n-1),14} & \cdots & 1 \end{pmatrix}, \tag{18}$$

where ρ are the correlation coefficients [13] between pairs of data, where

$$\rho = \begin{cases} 1 \quad \text{if} \quad i = k \quad \text{and} \quad j = l, \\ \rho_{i,jl} \quad \text{if} \quad i = k \quad \text{and} \quad j \neq l, \\ \rho_{j,ik} \quad \text{if} \quad i = k \quad \text{and} \quad j = l, \\ 0 \quad \text{otherwise.} \end{cases} \tag{19}$$

3.8 Ego Networks

The dataset includes connections between the CDO, its primary initial purchasers and CMF. Hence, the bipartite networks to be analysed are defined here as the ego networks (see Section 3.4) of the purchaser and the CMF. The 1-step neighbourhood of the CMF is also included as this shows the range of expertise the firm may or may not have based on the degree of the node and the attributes of its alters. The projections of the bipartite networks are converted into adjacency matrices as specified in Eq. 5 for use in ERGM and QAP.

4 Results and Analysis

4.1 Best Performing and Worst Performing CDOs

Using the rating system described in Section 3.2, the performance ratings are calculated for each CDO (an example is given in Tab. 3) and plotted for each reported date (see Fig. 5). The logarithm of the average downgrade rate is then calculated from the average rate between each consecutive downgrade event. Using the difference in performance rating between two consecutive downgrade events, divided by the difference in Unix time stamp. Its logarithm is then used to scale the rates. A Unix time stamp corresponds to the number of seconds that have passed since 1st January 1970. It is used to convert the date of each downgrade event into seconds so that the rate of downgrade can be determined. The final results are shown in Tab. 4.

The results of the ratings show a clear difference in how the CDOs performed, some starting to downgrade from as early as 2007. Fig. 6 shows that the 'top 5' performers had relatively slow average downgrade rates, meaning that the CMFs that managed these CDOs were some of the last to fall victim to the crisis. In stark contrast to this, Fig. 7 shows very fast average downgrade rates, seen by almost vertical lines. Also, the initial performance ratings of some of these CDOs were higher than those of the top performers, meaning their tranche values had a higher average credit rating (in some cases mostly Aaa ratings). In the crisis it is fair to say that these had a much higher drop in quality compared to the CDOs from Fig. 6.

4.2 Credit Rating Agency Experience

Most of the careers in the dataset are in the financial sector, made up of roles in investment banking, asset management and credit ratings, although other types of career paths include law practices, academia, etc. The latter types are grouped into one career state called Misc. The chronogram shown in Fig. 8 depicts the frequencies of career states at different positions in time. The three most frequent careers are investment banking, collateral management and asset management.

Cluster analysis applied to the data distinguishes four groups of trajectories through use of agglomerative hierarchical clustering, displayed as a dendrogram in Fig. 9. The clusters are represented in separate individual plots in Fig. 10, and the corresponding state frequency distribution plot in Fig. 11. All clusters appear homogeneous with regard to the nature of their states, their length and order, and the length of whole careers. Heterogeneity describes the state of diversity in the sequences. This is overall quite low due to the presence of so many instances of investment banking and collateral management. However, there is a clear distinction between the Type 2 and Type 3 clusters, which can be seen in Fig. 11, with asset management having a relatively high frequency in the Type 2 cluster and credit rating agencies having a relatively high frequency in Type 3 cluster.

The results provide evidence that there are some similarities in the sequences, as this is how optimal matching produces the clusters in a hierarchy. Particularly, clusters of the Type 2 and Type 3 dissociate between individuals who had previous careers in credit rating agencies before the Financial Crisis and those who had previous careers in asset management firms before the Financial Crisis. Since the sequences are ordered in terms of which CMF individuals worked for, it is likely that the Type 2 and Type 3 clusters have sequences of people from several distinct CMFs. Thus, using the performance rating system to measure how CDO characteristics of these sequences performed would test if there is a correlation between experience within asset management/credit rating agencies and CDO performance. However, with the sequence analysis method used here it is not possible to class sequences in terms of the matching CDO.

4.3 Shared Careers

For this objective, the results from the ERG model are used to show if there is evidence of shared careers in the networks. The probability estimates show the chance of a particular structure being present in a randomly generated network. If the probability is relatively high, this means that there is evidence to show that a structure is not there by chance and can explain the connotations this may have, provided that there is a clear difference between the structures of the best performing CDOs and the worst performing CDOs.

Firstly, the ego networks are generated for each CMF and investment bank included in Tab. 4. For example, the networks of MFS Investment and Greywolf are shown in Fig. 12 and Fig. 13 respectively, each with the connections to their respective investment banks, Merrill Lynch and Goldman Sachs. Visual inspection of Fig. 13 shows that there are a significant number of actor nodes that are connected to both Greywolf and Goldman Sachs (the CMF and investment bank), meaning these are people that have been employed by both firms at some point in their career. These types of patterns emerge more in the ego networks of the worse performing firms compared to the good performing firms. This can be seen when the ego networks of all the ten performers in Tab. 4 are combined. These networks are shown in Fig. 14 and Fig. 15. The degree distributions for nodes of the actor type are given in each. In Fig. 14, the agents are mostly of degree 1 (89%), meaning there is a low chance that an agent in the network has been employed by both a CMF and an investment bank that are both involved in the same CDO. However, there is a rise in higher degree actor nodes in Fig. 15, as the degree distribution shows that only 55.33% are of degree 1. This means that there is a higher chance that someone in the network was employed by both a CMF and an investment bank that are both involved in the same CDO.

The ERGM is now carried out on the projection of each generated ego network. The results of the simulations are shown in Tab. 5 and Tab. 6. An example of a trace and kernel density plot is shown in Fig. 16. The purpose of this is to interpret the performance of the MCMC sampler. The trace plot looks as it should as there are no flat sections, which would suggest that the chain stays in the state for too long. The kernel density plot shows no signs of major skewness, implying the posterior draws are accurate. The performance plots are tested for each simulation and the plots all have similar shapes to those in Fig. 16, meaning the results are reliable. The standard errors are also recorded in the tables and the largest error has size 0.1949. Thus, it is acceptable to say that the results are not hugely inaccurate in any way. Seneca Capital appears in both lists so the same estimates are recorded in each table.

When the results of a sample estimate are close to zero, the probability converges to 0.5 (see Fig. 4). A maximum likelihood estimate is generally considered statistically different from zero (i.e. statistically significant) if the confidence interval does not contain zero, or if the ratio of the estimate to the standard error exceeds 1.96 in magnitude [14]. Calculating the ratios for each CMF in Tab. 5, the levels of significance for each effect is as follows:

- Five out of five of the edge estimates are significant
- Five out of five of the triangle estimates are significant
- Three out of five of the two-star estimates are significant
- Two out of five of the three-star estimates are significant

On the other hand, calculating the ratios for each CMF in Tab. 6, the levels of significance for each effect is as follows:

- Five out of five of the edge estimates are significant
- Four out of five of the triangle estimates are significant
- Four out of five of the two-star estimates are significant
- Three out of five of the three-star estimates are significant

From these results, all of the edge estimates are significant. This is common in most undirected networks with several edges, as the feature is usually dominant in a randomly generated network. The most notable result is the number of significant two-star and three-star estimates. There are a higher number of both in Tab. 6, which suggests that two-star and three-star structures are more present in the observed networks of the bad performing CDOs than in the observed networks of the good performing CDOs. At least, this means that in a randomly generated network the probability of such structures appearing would be higher. However, for k-stars where k>3, the trend seems to show that these become less common as k increases. This is a reasonable assumption because it is unrealistic that an individual would have a shared career with more than 3 people in a small ego network like the ones simulated here, as people generally have very diverse careers.

As previously stated, to prove statistical significance the results generated by QAP have been tabulated in Tab. 7, 8, 9 and 10. Tab. 7 and 8 are tables of p-values for the pairs of CMFs in the networks, each for the top and bottom 5 performing CDOs. p-values here corresponds to the metric of statistical significance.

Tab. 9 and 10 correspond to the auto-correlation matrix. They depict the correlation coefficient for pairs of CMF in each of the top and worst 5 performing CDOs. By comparing the coefficients in both tables, the main thing that can be inferred. is that in the network of the 5 best performing CDO there are more negative correlations. Negative correlation implies an inverse behaviour between two variables: as one increases the other decreases proportionally to the correlation coefficient. For example, Vertical Capital and MFS Investment have a negative correlation coefficient of -0.081, this implies that as the ratio of number of employees present in both CMF increases, the less likely they will share common employees which shows a better case of homophily in the good performing networks.

4.4 Career Length at Firms

The results for the weighted degree distributions of the simulated ego networks are given in Fig. 17 and Fig. 18. The networks are first projected from bipartite into unipartite before the weighted degrees are plotted in these distributions. The distributions are all positively skewed, meaning most of the distribution is concentrated towards a degree of 0. This means that the majority of the nodes have a weight of zero. Since the weights translate as the average number of years at a firm shared between individuals, the skewness shows that there are a fair amount of individuals who worked for a firm independently of other people.

The variance in the distributions is higher for networks related to good performance, as in Fig. 17 there are more bars on average for the graphs than in Fig. 18. Also, Fig. 18d shows that there are 9 nodes in total and 2 of these have a weighted degree of 3, making up a proportion of 0.22 of the distribution. The highest weighted degrees have much lower proportions in Fig. 17, such as Fig. 17a which shows that only 1 out of 35 nodes has a weighted degree of 28. These results suggest that networks like those in Fig. 18d have a high proportion of people who had a relatively long career (compared to the rest of the population of nodes) at a particular firm, whereas distributions such as Fig. 17a suggest careers mostly in the range of 1-10 were shared. However, nodes with weight 0 in these distributions have the highest proportion.

5 Discussion and Conclusion

From the chronogram in Fig. 8, it is seen that careers gradually drifted away from investment banking - that is, for the individuals included in the dataset - and the collateral manager occupation type had a sharp increase in 'popularity' from about 2005 to 2008. Roughly 75% of careers in the dataset are in this category, showing evidence that the career progression was a result of collateral management firms being a very prosperous and enticing career option in the years leading up to the collapse, giving some indication of the widespread infiltration of the CDO during the course of the GFC.

The weighting of the nodes in the networks provides some indication of how long individuals worked at a particular firm. However, it does not account for the fact that at a particular firm the connected actors do not always work at the same firm at the same time. For example, an individual could have worked at an investment bank from 2004-2006 and a separate individual worked at the same firm from 2007-2008, hence there is no overlap in career duration at the same firm, yet they are connected in the bipartite network projection. The results produced by the ERG model shows promises of there being a difference in the number of k-stars between CMF ego networks that had better performing CDOs and CMF ego networks that had worse performing CDOs. However, there simply are not enough samples to determine that the obtained results can be considered concrete evidence. If there were more CMF networks simulated from the data of more CDOs, then it might be possible to use the model to show that the results are indeed significant. Unfortunately, ERGM cannot say to what extent a particular result is significant, but only gives a yes or no answer to whether it is significant or not.

Some of the CDOs performed irrespective of their collateral manager. For example, Seneca Capital managed two different CDOs, and by means of the rating system these appeared as both some of the highest and lowest performing CDOs. This is not true for all CMFs, because firms like Greywolf only managed one CDO so the results produced for these types of CDOs should be reflective of their performance during the GFC. However, it is important to note that companies like Merrill Lynch would have been involved in quite a few CDOs, making it difficult to distinguish whether the firm in question was responsible or if the issue was with individual malpractice.

5.1 Comments on the Model

There are further techniques in the area of sequence analysis that could be used to give more accurate results for the career sequencing method. For example, sequence mining enables the retreival of specifically defined sub-sequences, which would consider the credit rating agency state in more detail. As a whole, clusters resulting from a typology can be finely described by means of each cluster's individual graph, some sequence statistics, prototypical sequences and the list of frequent subsequences [3]. Bootstrapping and simulation would provide statistical analysis in a more detailed and accurate setting, in order to make the results more robust. Sequence analysis can also provide measures on the diversity of sequences. That is, individual sequences may be measured by their entropy, variance and the complexity of length. All of this could make an improvement to the sequencing method which may produce more reliable results and a deeper insight into how career progression could have an effect on CDO performance.

The issue with ERGM is the number of possible graphs that can be generated. As mentioned previously, the set of all possible networks has size $2^{n(n-1)/2}$, meaning the size of n must be kept as small as possible in order for the Markov Chain to fully converge. This convergence issue is detrimental to the results, as larger networks such as those in Fig. 14 and Fig. 15 would likely have interesting properties, but the attempts to produce these results ended in a failure to converge. This would result in 'diagnostic' plots that look significantly worse than Fig. 16. Further development and experimenting with the new specifications is required to attain a well-balanced methodology for the statistical modelling of social networks [15], especially in a setting of so many career paths. Higher order structural effects are also further improvements that could be added to the model. However, more variety of data would be needed to make these improvements, such as attributes of the employees (e.g. positions, projects, etc.). This is also true for QAP.

5.2 Further Steps

With regards to this research, there are a number of further avenues yet to be explored and further improvements that can be made to the analytic techniques presented in this report. Quantifying CDO performance could be significantly improved by employing survival analysis, a branch of statistics for analysing the expected duration of time until one or more events happen; in this case, credit rating downgrade events. Further analysis would also include consideration of career sequence analysis for individual CDOs or CMFs. This would allow one to determine whether specific patterns in career history are pertinent to the performance of individual CDOs or a group of CDOs managed by a particular group of actors. The final and likely most influential further step would the implementation of career 'overlap' as a method of edge weighting in the ERGM. Consideration of the number of years that individual actors have shared careers would be greatly beneficial in understanding the level of interaction between certain members of the career network.

References

[1] Moody's Asia Pacific Desktop Reference, "Moody's rating scale and definitions." https://www.moodys.com/sites/products/ProductAttachments/AP075378_1_1408_KI.pdf, (Accessed: 09/05/2019).

[2] Moody's Investment Service Compliance Department, "Policy of withdrawal of credit ratings." https://www.moodys.com/uploadpage/Mco%20Documents/SP13418_PolicyforWithdrawalofCR.pdf, (Accessed: 09/05/2019).

[3] P. Blanchard, "Sequence analysis for political science." APSA 2011 Annual Meeting Paper, 2011.

[4] E. Driver and L. Alfred, "Quantitative expression of cultural relationships." University of California Publications in American Archaeology and Ethnology. Quantitative Expression of Cultural Relationships: 211–256, 1932.

[5] A. Abbott and A. Tsay, "Sequence analysis and optimal matching methods in sociology: Review and prospect." Sociological Methods Research, 29(3-33), 2000.

[6] L. Lesnard, "Optimal matching and social sciences." halshs.archives-ouvertes.fr, 2006.

[7] J.H. Ward, "Hierarchical grouping to optimize an objective function." Journal of the American Statistical Association, 58, 236–244, 1963.

[8] Y. Fan, M. Li, P. Zhang, J. Wu and Z. Di, "The effect of weight on community structure of networks." PHYSICA A 378: 583-590, 2007.

[9] D. Hunter, M. Handcock and M. Morris, "ergm: A package to fit, simulate and diagnose exponential-family models for networks." Journal of Statistical Software, 24(3), 2008.

[10] C. Bishop, "Pattern recognition and machine learning." Springer, ISBN 978-0-387-31073-2, 2006.

[11] T. Snijders, P. Pattison, G. Robins and M. Handcock, "New specifications for exponential random graph models." Sociological Methodology, 36(1), 99-153, 2006.

[12] T. Snijders, "Markov chain monte carlo estimation of exponential random graph models." ICS, Department of Statistics and Measurement Theory University of Groningen, 2002.

[13] E. Lawler, "The quadratic assignment problem." Management Science, 9(4), 517-708, 1963.

[14] C. Schmid and B. Desmarais, "Exponential random graph models with big networks: Maximum pseudo-likelihood estimation and the parametric bootstrap." 2017 IEEE International Conference on Big Data (Big Data), Boston, MA, USA, 2017.

[15] G. Robins, T. Snijders, P. Wang, M. Handcock and P. Pattison, "Recent developments in exponential random graph (p*) models for social networks." Social Networks 29, 192–215, 2007.

[16] D. Vaidya, "Dividend discount model." https://www.wallstreetmojo.com/asset-backed-securities/, (Accessed: 09/05/2019).

Appendix

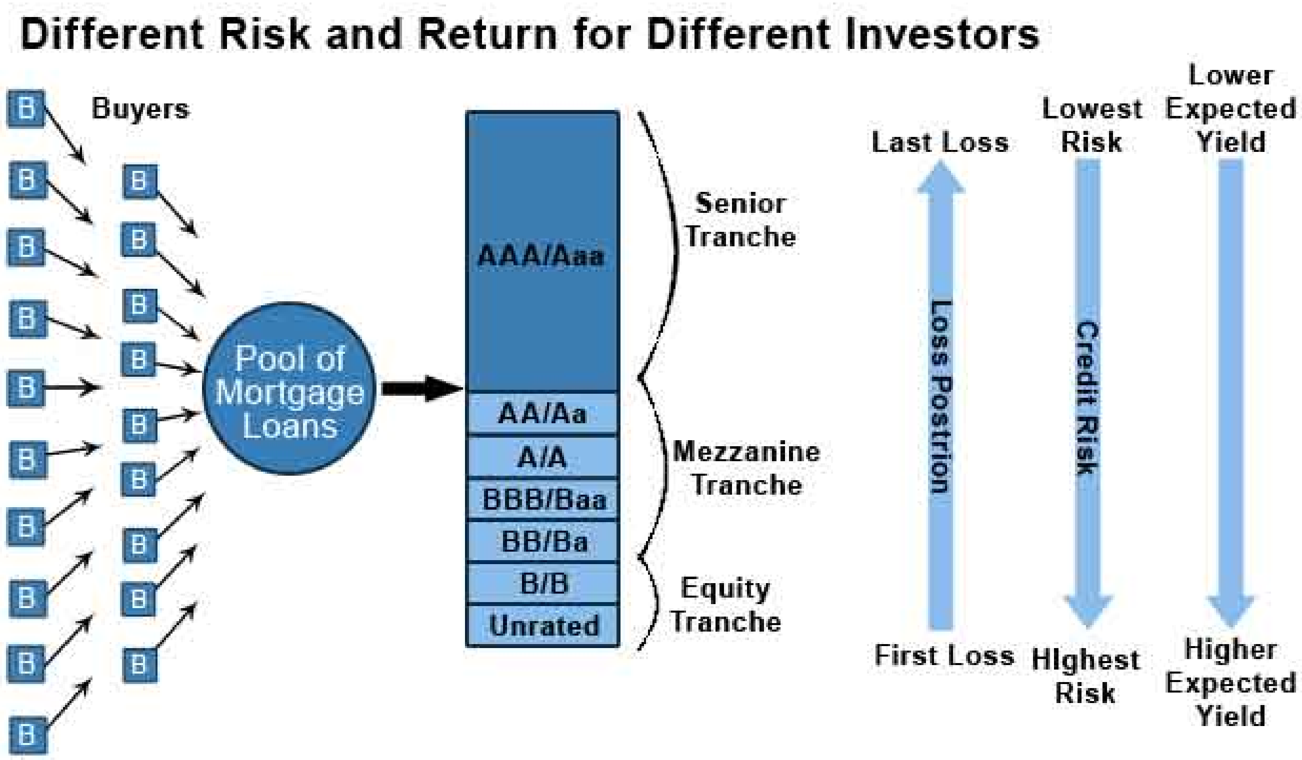

Figure 1: Diagram detailing the creation and structure of a CDO [16] with the respective risk and return profiles.

Calculation of Numerical Rating	
Moody's Rating	**Numerical Rating Score**
Aaa	21
Aa1	20
Aa2	19
Aa3	18
A1	17
A2	16
A3	15
Baa1	14
Baa2	13
Baa3	12
Ba1	11
Ba2	10
Ba3	9
B1	8
B2	7
B3	6
Caa1	5
Caa2	4
Caa3	3
Ca	2
C	1
W	0

Table 1: Above is the method used for converting letter-grade ratings into a numerical score for use in calculating the rate of downgrade.

	Downgrade Event						
	30/03/07	**31/10/07**	**28/02/08**	**09/05/08**	**23/09/08**	**14/07/09**	**17/07/09**
Tranche Rank (From Highest Value to Lowest Value)	**Moody's Rating**						
1	Aaa	Aaa	Ba1	B3	B3	Ca	W
2	Aaa	A2	Caa3	Ca	C		
3	Aaa	A3	Caa3	Ca	C		
4	Aa2	Baa2	Ca	Ca	C		
5	Aa3	Baa3	C				
6	A2	Ba2	C				
7	A3	Ba3	C				
8	Baa2	Caa2	C				
9	Baa3	Ca	C				
10	Baa3	C					
11	Ba2	C					

Table 2: An example of the reported downgrade events for a CDO. This is the Auriga 2007 CDO which is composed of 11 tranches. The Moody's Rating for each tranche is given.

	Downgrade Event						
	30/03/07	**31/10/07**	**28/02/08**	**09/05/08**	**23/09/08**	**14/07/09**	**17/07/09**
Tranche Rank (From Highest Value to Lowest Value)	**Moody's Rating**						
1	21	21	11	6	6	2	0
2	21	16	3	2	1		
3	21	15	3	2	1		
4	19	13	2	2	1		
5	18	12	1				
6	16	10	1				
7	15	9	1				
8	13	4	1				
9	12	2	1				
10	12	1					
11	10	1					
Total	**178**	**104**	**24**	**12**	**9**	**2**	**0**

Table 3: Table showing the ratings of the Auriga 2007 CDO converted into scores using the system from Tab. 1. The totals of the scores are also given.

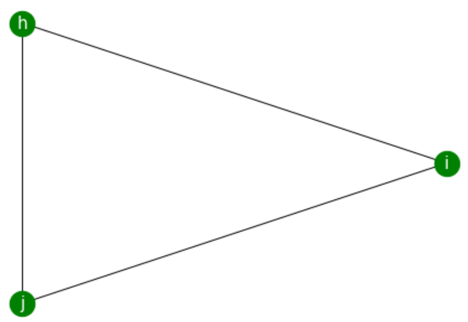

Figure 3a: A triangle structure which has the triadic closure property between actors/nodes i, j and h.

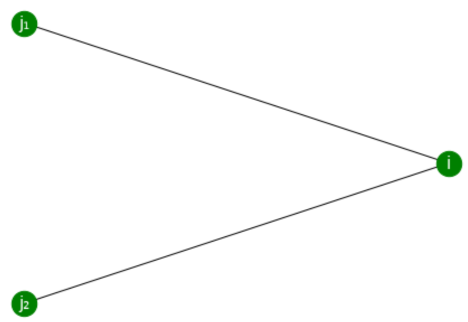

Figure 3b: A two-star structure with central node i and its neighbours j_1 and j_2.

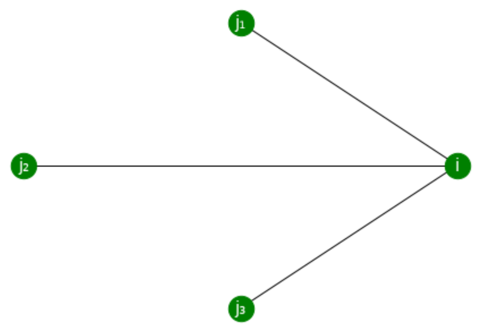

Figure 3c: A three-star structure with central node i and its neighbours j_1, j_2 and j_3.

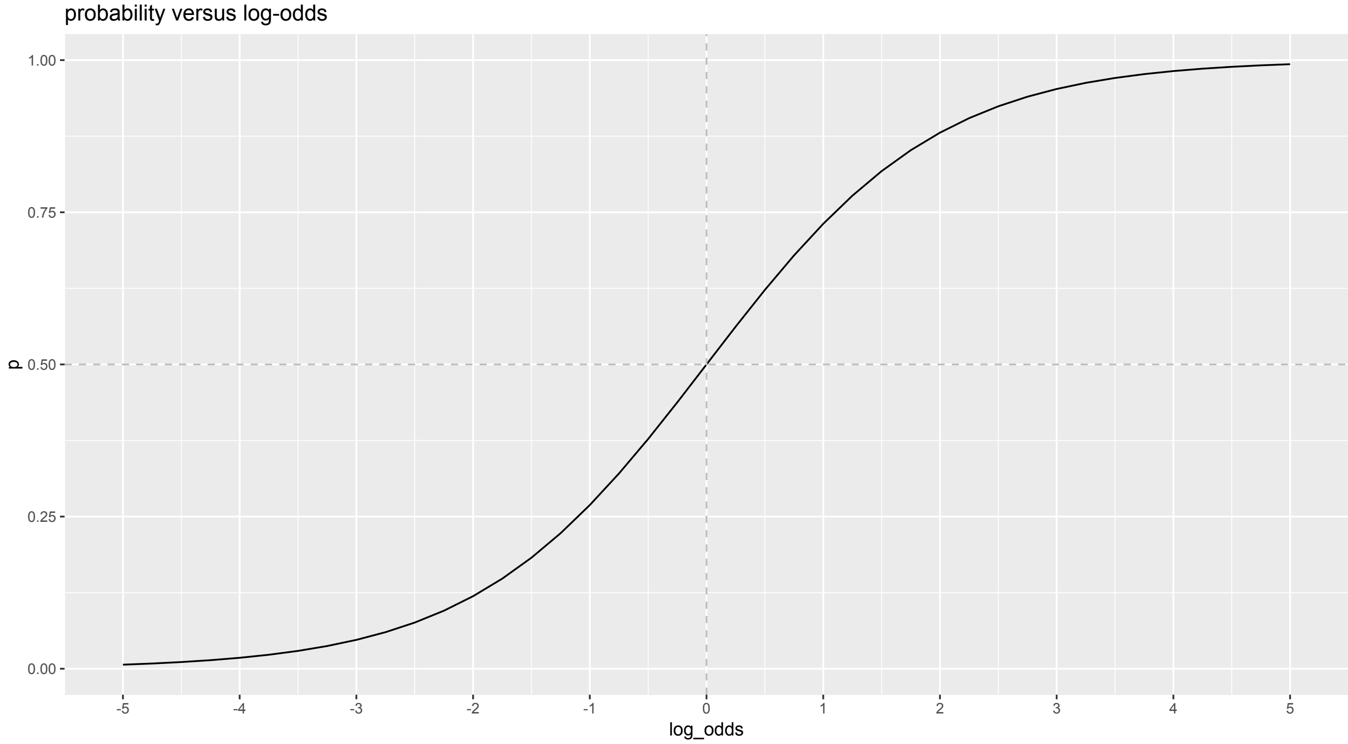

Figure 4: Plot of the range of log-odds estimate against the probability of tie formation. This is used to convert the estimation from the MCMC into probabilities.

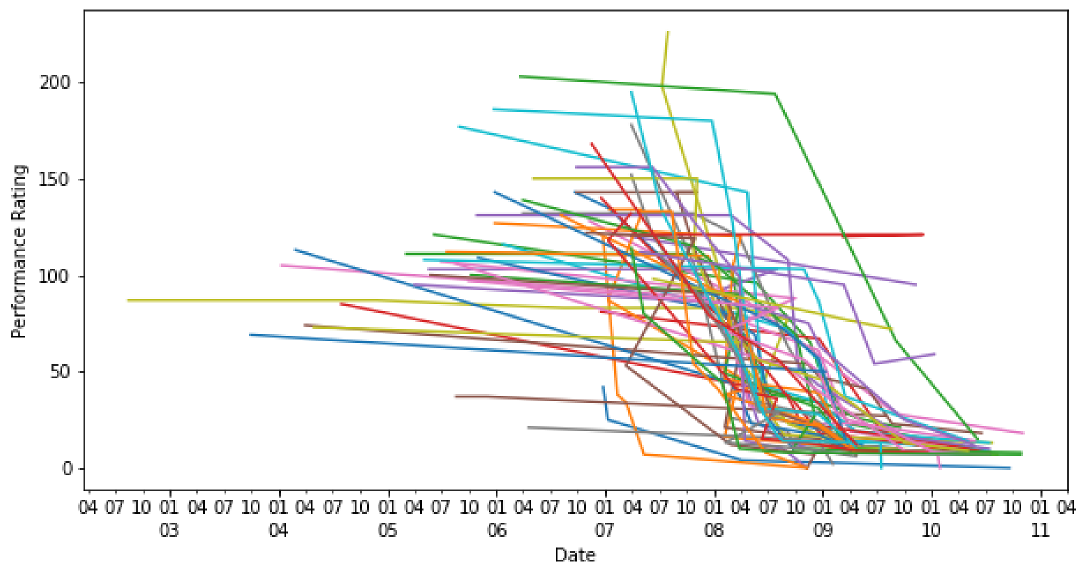

Figure 5: Graph depicting the downgrading of all 55 CDOs. Performance Rating is calculated from Tab. 3.2 and the start and end dates (given in month and year) of each CDO stems from their first and final reported dates.

Rank	CDO Name	Average Log Downgrade Rate	CMF	Bank
1	Newbury Street	-11.54	MFS Investment	Merrill Lynch
2	Pampelonne I	-11.91	Vertical Capital	Barclays
3	Fort Sheridan ABS	-12.09	Vanderbilt Capital	Merrill Lynch
4	Caldecott I	-12.15	Seneca Capital	Merrill Lynch
5	GSC ABS	-12.16	GSC Partners	Merrill Lynch
⋮	⋮	⋮	⋮	⋮
51	Broderick II	-15.01	Seneca Capital	Merrill Lynch
52	Duke Funding III	-15.53	Duke Capital	Wachovia Securities
53	Adams Square II	-15.99	Credit Suisse	Citigroup
54	Octans I	-16.96	Harding Advisory	Merrill Lynch
55	Timberwolf I	-18.06	Greywolf Capital	Goldman Sachs

Table 4: Table showing the calculated downgrade rates of each CDO. The rate is calculated from the average rate between each consecutive downgrade event (using the difference in rating divided by the difference in Unix time stamp) and its logarithm is then used to scale the rates. Highlighted in green are the 5 CDOs that have the lowest magnitude average downgrade rate. Highlighted in red are the 5 CDOs that have the highest magnitude downgrade rate. Also given are the CMFs who managed the CDO and the bank who purchased the CDO.

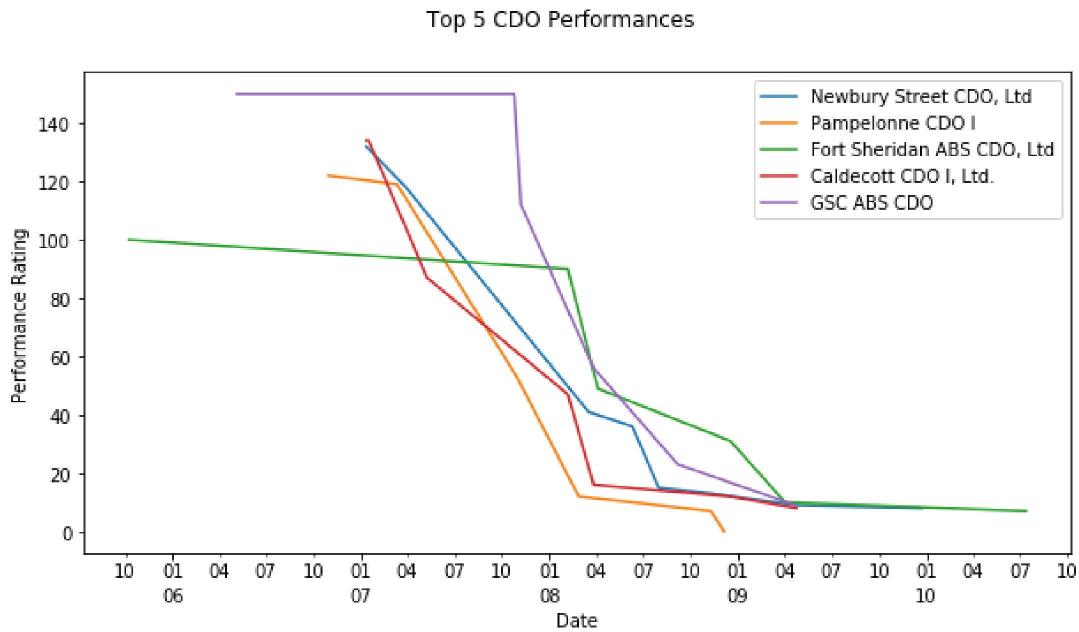

Figure 6: Filtered version of Fig. 5 that only includes the 5 CDOs highlighted in green in Tab. 4.

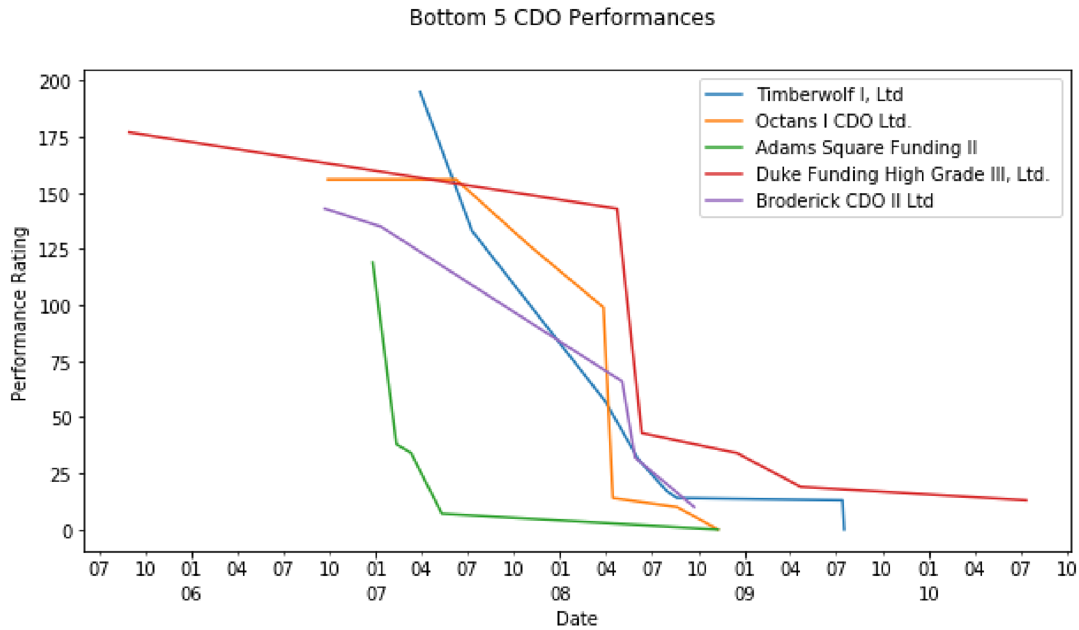

Figure 7: Filtered version of Fig. 5 that only includes the 5 CDOs highlighted in red in Tab. 4.

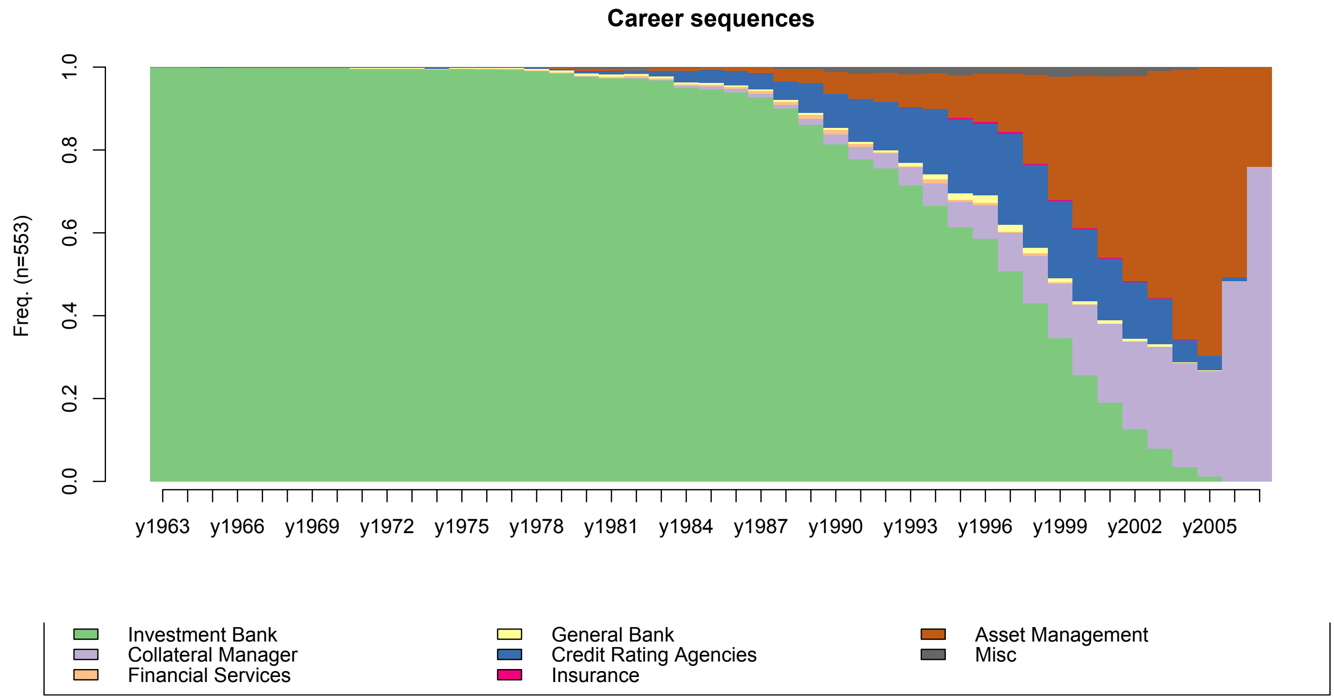

Figure 8: State distribution plot (chronogram) showing the state frequency of each career type at a position of time from 1963 to 2008. By the start of 2008, about 75% of all individuals have transitioned into working for a collateral manager.

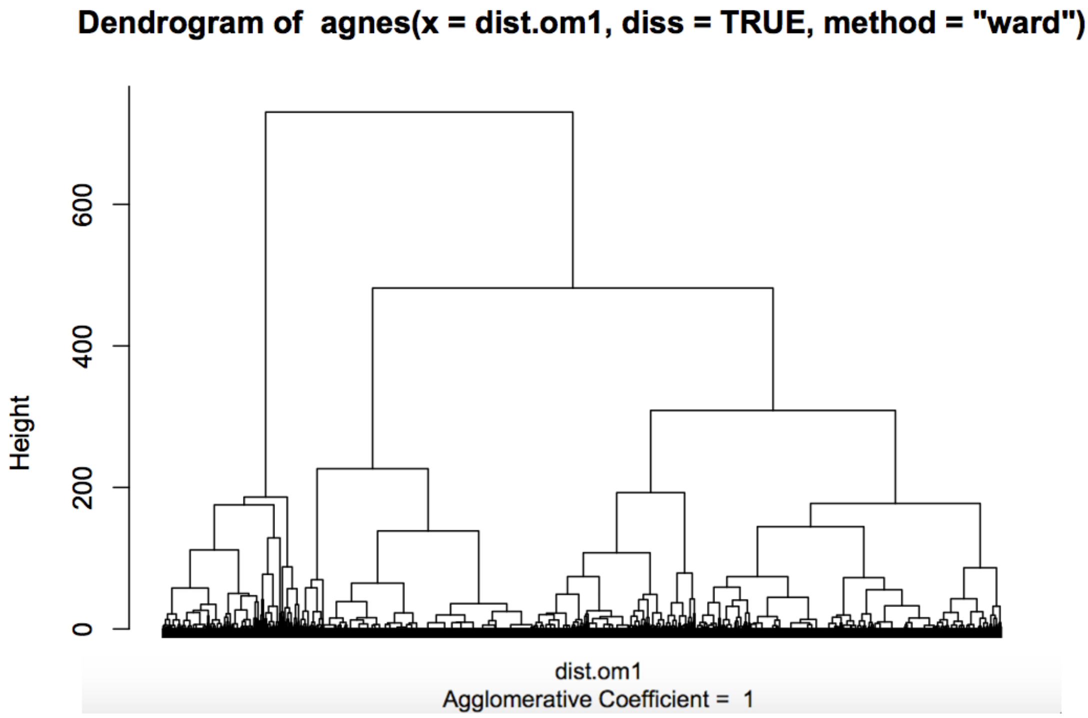

Figure 9: Dendrogram of optimal matching analysis, found by using Ward's method to build a hierarchy of clusters. Agglomerative clustering refers to each observation starting in its own cluster, and pairs of clusters being merged as one moves up the hierarchy.

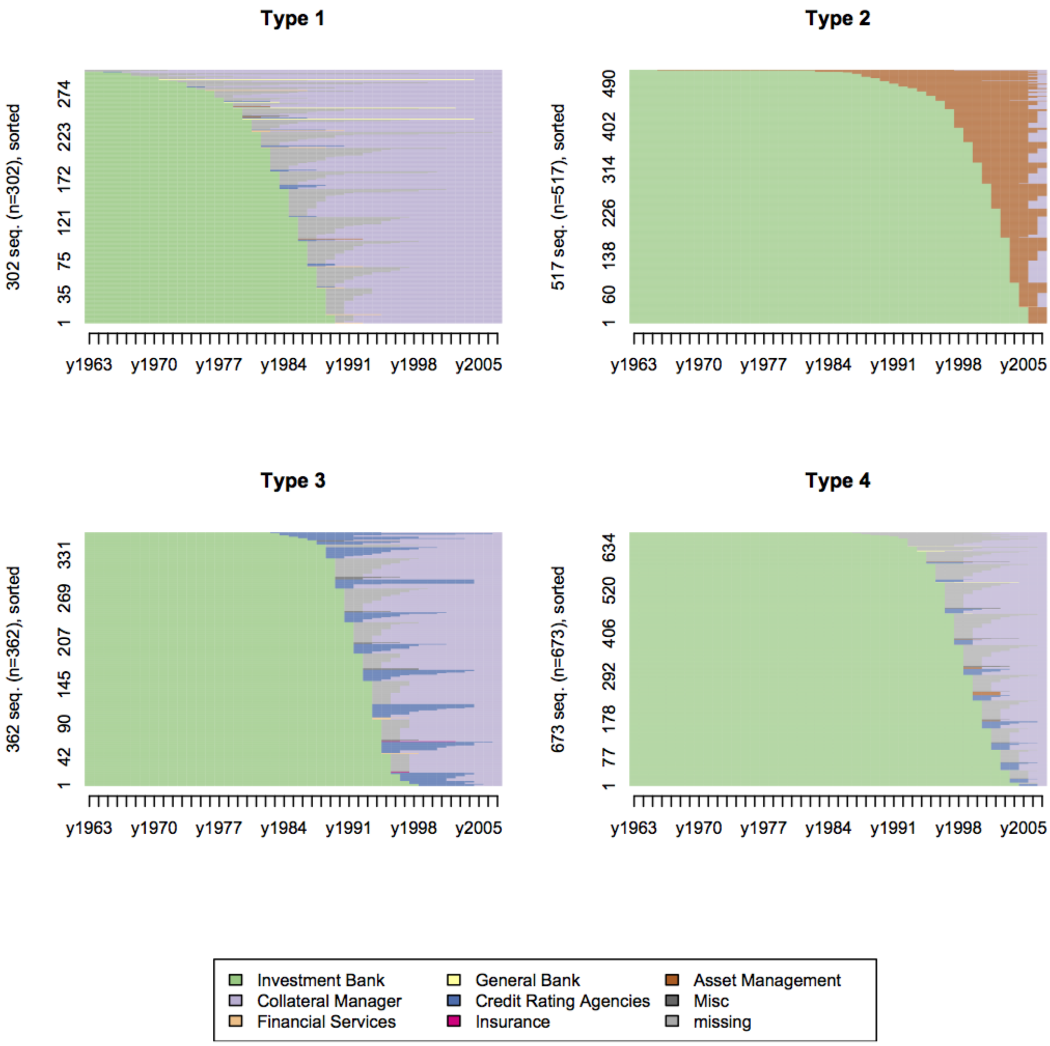

Figure 10: Sequence index plot showing individual sequences sorted as stacked bars for each of the four different cluster types grouped by hierarchical clustering.

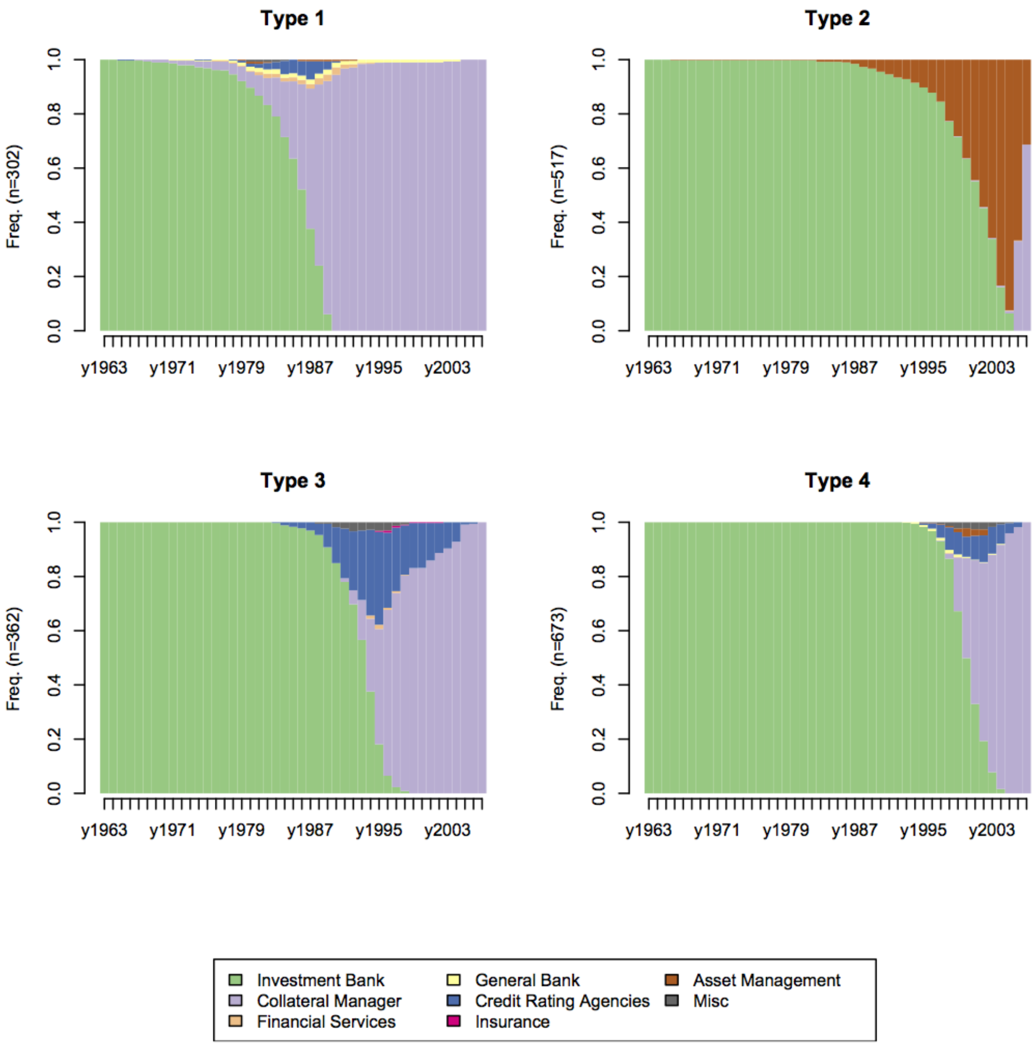

Figure 11: State distribution plot showing the state frequency of the four different cluster types grouped by hierarchical clustering.

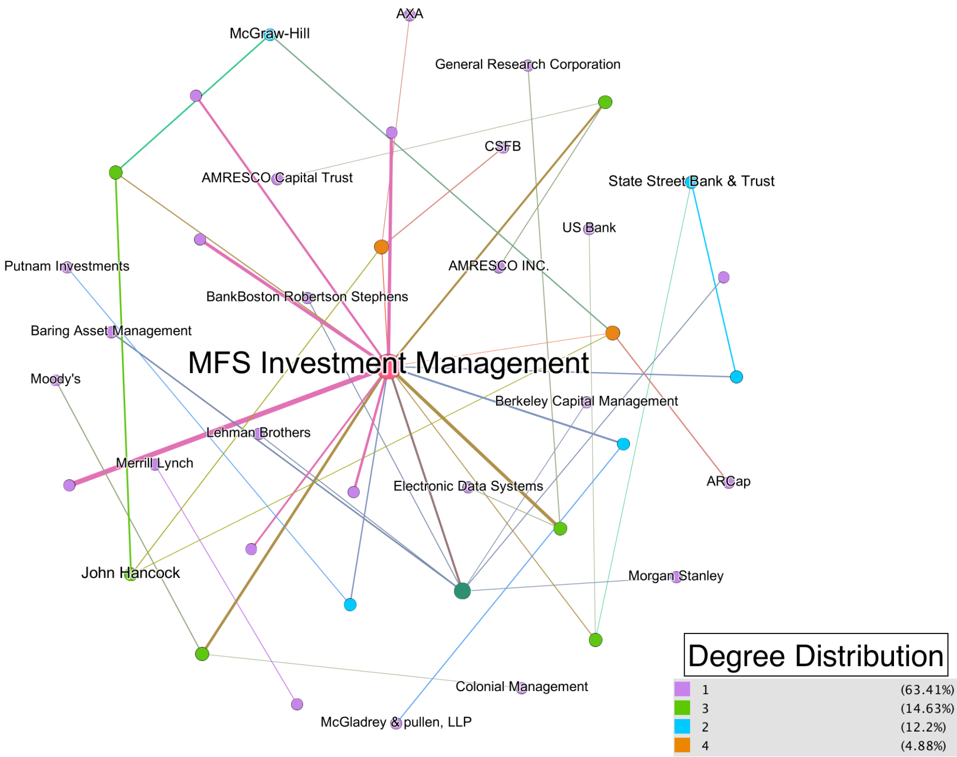

Figure 12: Example of an ego network with depth 2 for MFS Investment, a top performing CMF. Also included is the bank purchaser, Merrill Lynch, with depth 1.

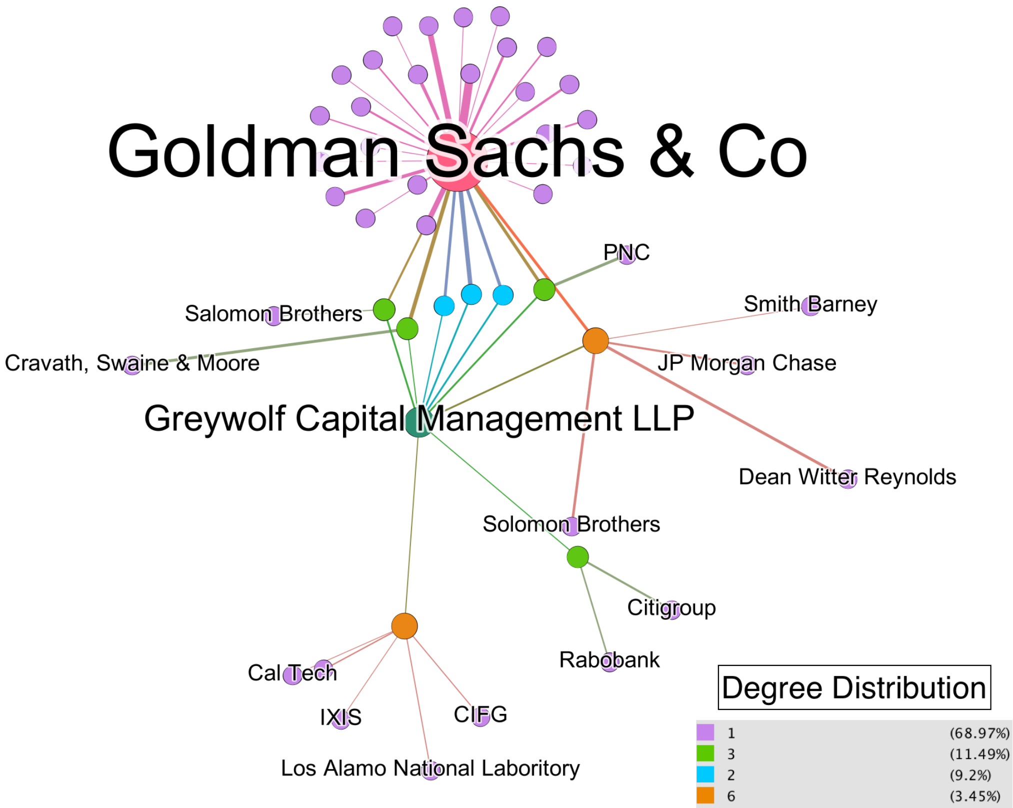

Figure 13: Example of an ego network with depth 2 for Greywolf Capital Management, a bottom performing CMF. Also included is the bank purchaser, Goldman Sachs, with depth 1.

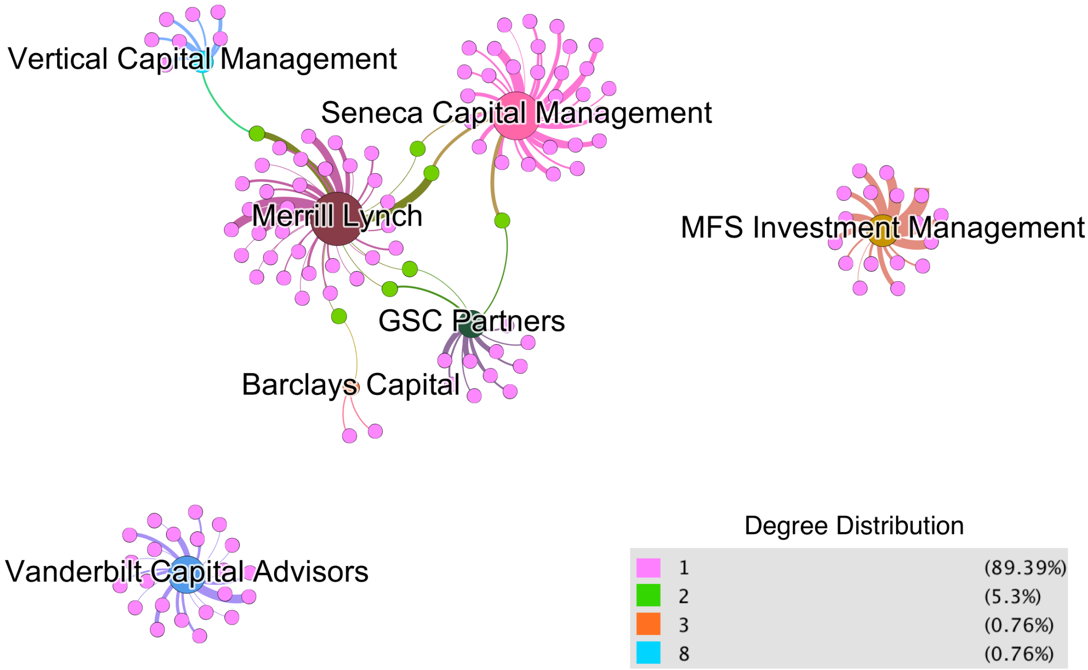

Figure 14: Combined ego networks of depth 1 for each of the top performing CMFs and their corresponding bank purchasers.

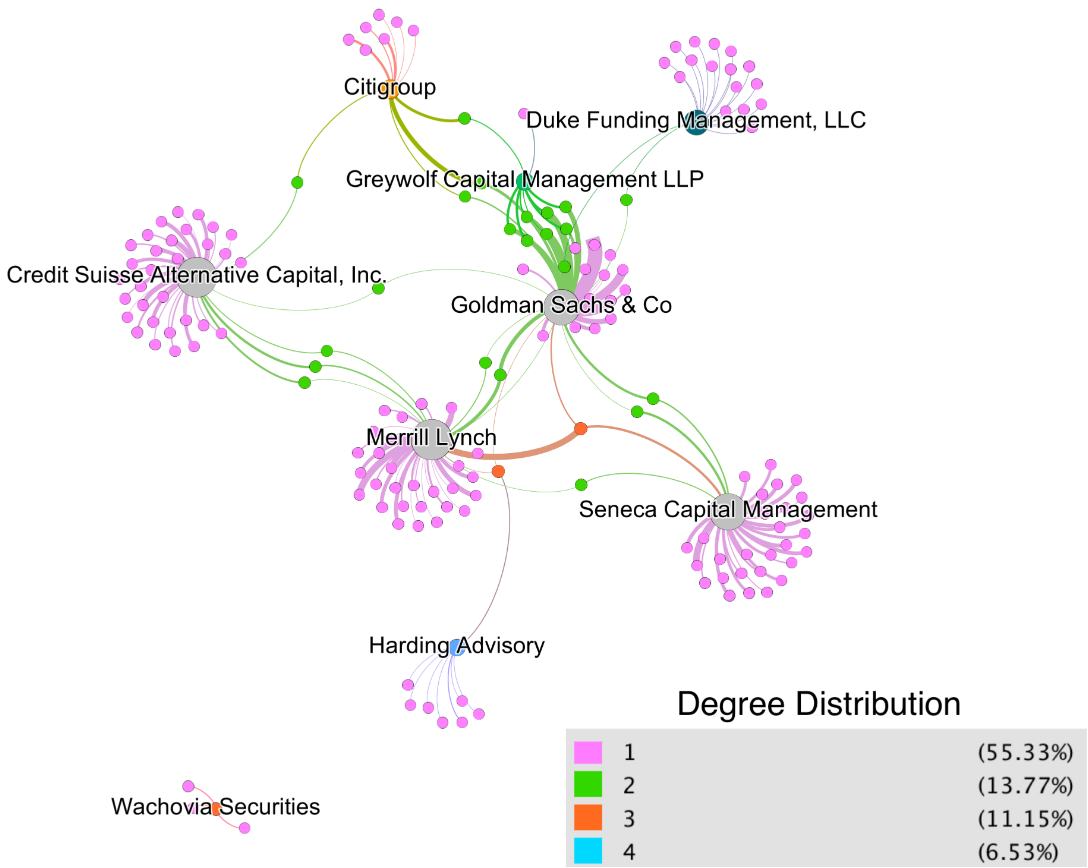

Figure 15: Combined ego networks of depth 1 for each of the bottom performing CMFs and their corresponding bank purchasers.

CMF Name	Effect	Log-odds Estimate	Probability, p(y)	Standard Error
MFS Investment	Edges	-0.07691	0.4808	0.05291
	Triangles	0.022168	0.5055	0.002275
	Two-stars	0.0001558	0.5000	0.0009878
	Three-stars	1.184e-04	0.5000	6.997e-05
Vertical Capital	Edges	-1.4534	0.1895	0.1851
	Triangles	-0.4762	0.3831	0.1213
	Two-stars	-0.1219	0.4696	0.0240
	Three-stars	-0.01759	0.4956	0.00486
Vanderbilt Capital	Edges	-2.2407	0.0962	0.1215
	Triangles	-0.86524	0.2962	0.08941
	Two-stars	-0.12257	0.4694	0.01123
	Three-stars	-0.0076133	0.4982	0.0009818
Seneca Capital	Edges	-0.09143	0.4772	0.04258
	Triangles	0.015988	0.5039	0.001595
	Two-stars	-0.0008325	0.4998	0.0006878
	Three-stars	-9.280e-06	0.5000	4.131e-05
GSC Partners	Edges	0.42100	0.6037	0.06087
	Triangles	0.037471	0.5094	0.002011
	Two-stars	0.0094261	0.5024	0.0009443
	Three-stars	7.265e-04	0.5002	5.184e-05

Table 5: Results of the ERG model for the 5 best performing CMF ego networks. Log-odds estimates are converted into probability of tie formation, $p(y)$, according to the plot in Fig. 4.

CMF Name	Effect	Log-odds Estimate	Probability, p(y)	Standard Error
Seneca Capital	Edges	-0.09143	0.4772	0.04258
	Triangles	0.015988	0.5039	0.001595
	Two-stars	-0.0008325	0.4998	0.0006878
	Three-stars	-9.280e-06	0.5000	4.131e-05
Duke Capital	Edges	1.036	0.7381	0.157
	Triangles	-0.09182	0.4771	0.04478
	Two-stars	0.043250	0.5108	0.004776
	Three-stars	0.0063414	0.5016	0.0005449
Credit Suisse	Edges	-2.46640	0.0782	0.07365
	Triangles	-0.69887	0.3321	0.04488
	Two-stars	-0.096197	0.4760	0.004797
	Three-stars	-0.0008861	0.4998	0.0000740
Harding Advisory	Edges	-0.9293	0.2831	0.1142
	Triangles	-0.11055	0.4724	0.03081
	Two-stars	-0.040001	0.4900	0.007433
	Three-stars	-0.0032399	0.4992	0.0009154
Greywolf Capital	Edges	0.042	0.3250	0.1949
	Triangles	-0.07413	0.4815	0.05463
	Two-stars	-0.05063	0.4873	0.02038
	Three-stars	-0.005489	0.4986	0.004308

Table 6: Results of the ERG model for the 5 worst performing CMF ego networks. Log-odds estimates are converted into probability of tie formation, $p(y)$, according to the plot in Fig. 4.

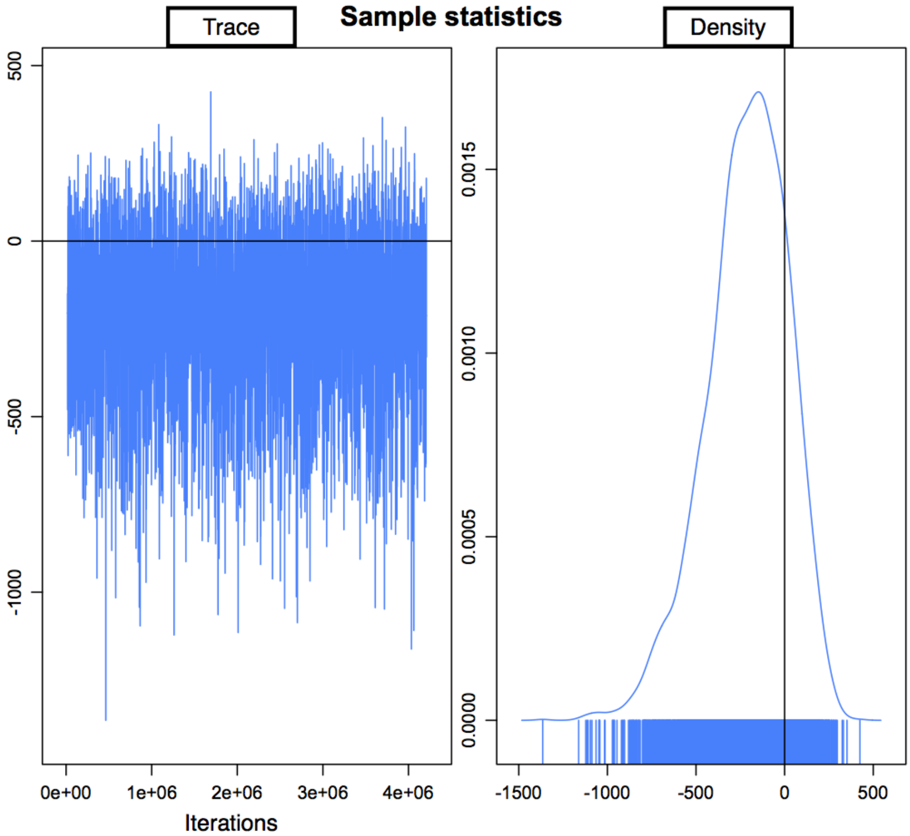

Figure 16: On the left is the trace plot for a typical MCMC sampling method showing the number of iterations the chain stays in each state. On the right is the posterior distribution density plot.

	MFS Investment	Vertical Capital	Vanderbilt Capital	Seneca Capital	GSC P.
MFS Investment	0	0.039	0.517	0.237	0.031
Vertical Capital	0.039	0	-0.079	0.467	0.061
Vanderbilt Capital	0.517	0.042	0.042	0.234	0.162
Seneca Capital	0.237	0.467	0.234	0	0.571
GSC Partners	0.031	0.061	0.162	0.571	0

Table 7: Matrix of p-values between dyadic connections (weights) for the 5 best performing CMF.

	Seneca Capital	Duke Capital	Credit Suisse	Harding Advisory	Greywolf Capital
Seneca Capital	0	0.025	0.032	0.055	0.016
Duke Capital	0.025	0	0.025	0.016	0.029
Credit Suisse	0.032	0.025	0	0.234	0.162
Harding Advisroy	0.055	0.016	0.234	0	0.571
Greywolf Capital	0.016	0.029	0.162	0.571	0

Table 8: Matrix of p-values for dyadic connections (weights) between the 5 worst performing CMF.

	MFS Investment	Vertical Capital	Vanderbilt Capital	Seneca Capital	GSC.P
MFS Investment	1	-0.081	-0.021	0.2	-0.073
Vertical Capital	-0.081	1	-0.007	-0.074	0.031
Vanderbilt Capital	-0.021	-0.007	1	-0.081	-0.05
Seneca Capital	0.2	-0.074	-0.081	1	-0.018
GSC Partners	-0.073	0.031	-0.05	-0.018	1

Table 9: Matrix of correlation coefficient between dyadic connections (weights) for the 5 best performing CMF.

	Seneca Capital	Duke Capital	Credit Suisse	Harding Advisory	Greywolf Capital
Seneca Capital	1	0.86	-0.6	-0.5	0.156
Duke Capital	0.86	1	0.057	-0.146	-0.079
Credit Suisse	-0.6	0.057	1	-0.076	0.7
Harding Advisory	-0.5	-0.146	-0.076	1	0.89
Greywolf Capital	0.156	-0.079	0.7	0.89	1

Table 10: Matrix of correlation coefficient between dyadic connections (weights) for the 5 worst performing CMF.

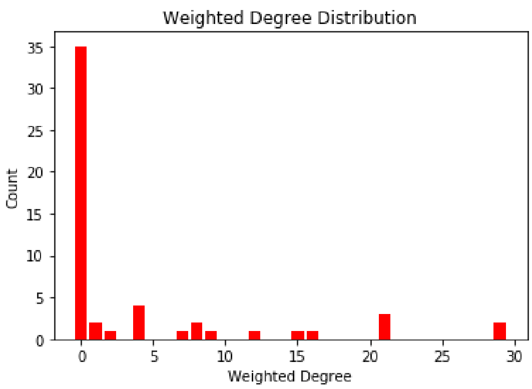

Figure 17a: Weighted degree distribution of the bipartite network for the MFS Investment ego network.

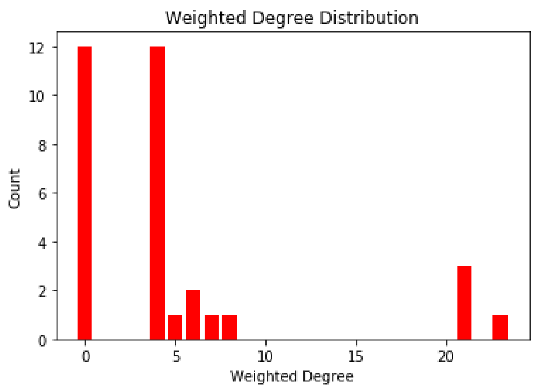

Figure 17b: Weighted degree distribution of the bipartite network for the Vertical Capital ego network.

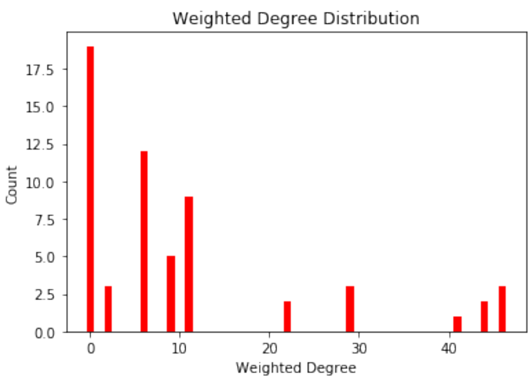

Figure 17c: Weighted degree distribution of the bipartite network for the Vanderbilt Capital ego network.

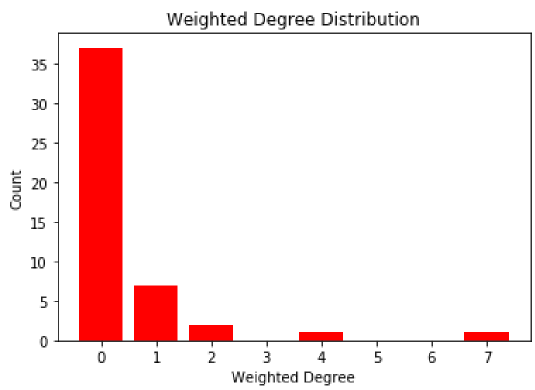

Figure 17d: Weighted degree distribution of the bipartite network for the GSC Partners ego network.

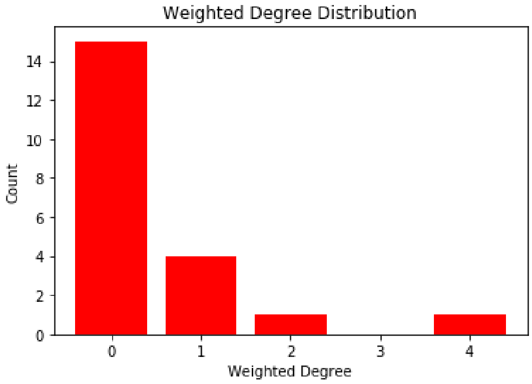

Figure 18a: Weighted degree distribution of the bipartite network for the Duke Capital ego network.

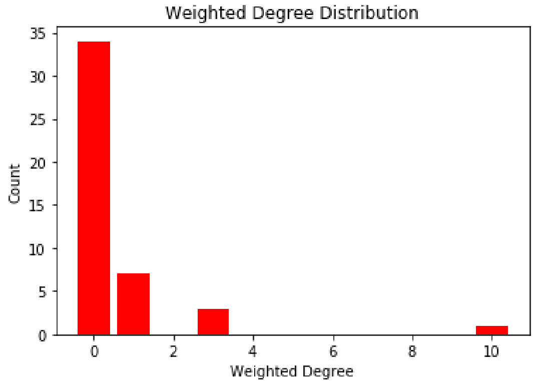

Figure 18b: Weighted degree distribution of the bipartite network for the Credit Suisse ego network.

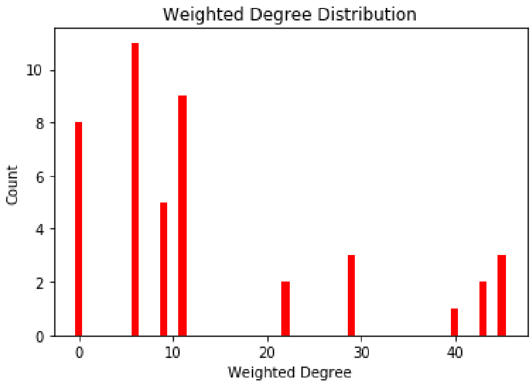

Figure 18c: Weighted degree distribution of the bipartite network for the Harding Capital ego network.

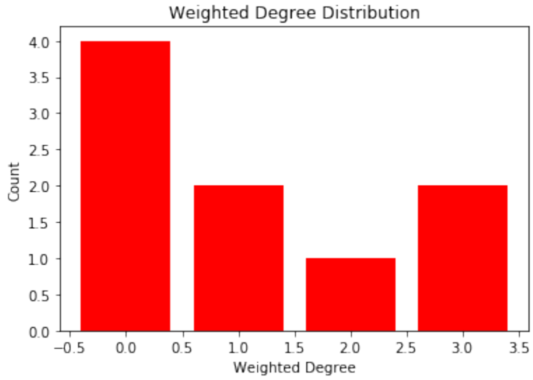

Figure 18d: Weighted degree distribution of the bipartite network for the Greywolf Capital ego network.