Irving Fisher Committee on
Central Bank Statistics

11th Biennial IFC Conference on “Post-pandemic landscape for central bank statistics”

BIS Basel, 25-26 August 2022

Examining concentration and similarity in institutional investors’ holdings[1]

Ariel Mantzura and Michael Bell,
Bank of Israel

[1] This presentation was prepared for the conference. The views expressed are those of the authors and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the event.

Examining Concentration and Similarity in Institutional Investor's Holdings

An application Using Network Analysis

Ariel Mantzura, Bank of Israel

Examining Concentration and Similarity in Institutional Investor's Holdings

The stock market is considered to be a good example of a complex system, one that consists of many entities and agents, interacting in such a way that their collective market behaviour is beyond a simple combination of their individual behaviours. Such systems are also characterized by various dimensions of interactions, that can be measured and modelled using data that is becoming increasingly more and more available. A key challenge in modern finance is finding efficient ways of summarizing and visualizing the stock market data to obtain useful information about its behaviour. Over the past two decades methodologies have been introduced to address such challenges, including machine learning techniques, and network analysis methods that have become very popular. In this study, we test how interdependence both between the agents and their assets within the network of portfolio holdings can be a source of systemic vulnerability. We study a real-world, institutional investors holdings network based on a granular dataset of all common assets (stocks and bonds) holdings in Israel from 2010-2021 and compare it with various alternative scenarios from randomization and rebalancing of the original investments. The scenarios generation relies on algorithms that satisfy the global constraints imposed by the numbers of outstanding shares in the market and the regulatory constraints. We extensively analyse the interplay between portfolio diversification and differentiation and examine how the outreach of exogenous shocks depends on these factors as well as on the type of shock and the size of the network with respect to the market. We find that real portfolios are diversified but highly similar, that portfolio similarity correlates with systemic fragility and that rebalancing can come with an increased similarity depending on the initial network configuration.

Keywords: Network analysis, cosine similarity, concentration

JEL classification: D85, G01, G18

Contents

1. Introduction

Finding efficient ways of summarizing and visualizing the stock market data to extract useful information about its behavior as well as finding interconnectedness patterns among market participants is crucial for identification of systemic vulnerabilities in the financial system and particularly, vulnerabilities arising from commonalities in asset holdings (overlapping portfolios). Using Network analysis for the Modelling of financial systems has become popular in recent years (Poledna et.al, 2020).

The aim of this study is to empirically analyse systemic vulnerabilities due to commonalities in asset holdings in the Israeli financial system across different types of financial institutions: Insurance companies, Mutual funds, Provident funds and Pension funds. The Institutional investors in Israel are required to publish on a quarterly basis their detailed asset holdings in accordance with the Capital Market Authority, Insurance and Savings reporting directives. Similarly, Mutual fund managers are required to publish on a monthly basis their detailed asset holdings in accordance with the Israeli Securities Authorities reporting directives. Using such required disclosures can allow for a study of commonalities across such investors.

For this study we use Praedicta's[1] Granular Data base of Management company asset holdings (equity and bonds) on a quarterly basis. Specifically, we test how interdependence both between the agents and their assets within the networks of portfolio holdings can be a source of systemic vulnerability. We study a real-world, institutional investors holdings network based on a granular dataset of all common assets (equity and bonds) holdings in Israel from 2010-2021.

Systemic vulnerabilities in financial markets may lead to the realization of systemic risks where exogenous shocks, failure or distress in one or several institutions may affect other financial institutions through some contagion channel. There are several contagion channels referred to in the literature. Poledna (2020) distinguishes between two types of contagion channels: direct and indirect contagion. Direct contagion (see Upper and Worms (2002)) is due to bilateral exposures of two institutions, i.e., through borrowing and lending and is quantified in the institutions balance sheet liabilities. Alternatively, indirect contagion can arise when financial institutions invest in the same assets where this type of contagion is referred to as overlapping portfolios (Levy-Carciente at al., 2015). In this case contagion has the following mechanism: As a result of some exogenous shock to some financial institution or to one or some of its assets it is forced to liquidate some of its securities at lower prices due to binding constraints i.e., contractual or regulatory constraints. Liquidation of securities at discounted prices referred to as fire sales propagates through the financial system hence causing losses to other financial institutions holdings those assets. Girardi et al. (2018) studied empirically this notion of indirect contagion in the insurance industry using the cosine similarity measure. Huang et al. (2013), Caccioli (2015) model overlapping portfolios as a bipartite network (banks and assets) of the banking system to describe the risk propagation process during crises and find that

[1] http://praedicta.com

their model can be useful for systemic risk stress testing for the banking system. Elsinger et al. (2006) examined systemic risk in the Austrian banking system and found that correlations in banks asset portfolios dominates contagion as the main source of systemic risk. Barucca et al. (2021) examined interconnections and systemic vulnerabilities related to price-mediated contagion (overlapping portfolios) across multiple types of institutions.

The paper is organized as follows: In section 2 we describe to more detail the data we use in our investigation and present in section 3 some descriptive statistics and various measures of concentration and similarity in asset holdings. In section 4 we model the financial system as a network and extract important measures.

2. Data

We use in this work Praedicta's granular Data base. The data base includes the total asset holdings of management companies (including institutional investors) from 2010Q1 up to the present2. The management companies are one of four: Insurance companies, pension funds, provident funds and mutual funds. The data is quarterly and detailed up to a particular asset and investment channel (funds). For each company and fund the total holding in assets is given. In Table 1 is an example of typical records of the data base.

Table 1: Example of Typical Records in Database

Management company Menora Mivtachim provident and pension			
Channel Name	Asset type	Asset name	Market Value
Fund A investment channel for age 60+	Corporate bond	Poalim bond 29	279,855
Fund A general investment channel	Equity TA 35	Elbit Systems	25,955,137
fund B	Corporate bond	Poalim bond 29	32,811
fund B	Equity TA 35	Elbit Systems	856,926
fund C	Corporate bond	Poalim bond 29	8,552,713
fund C	Equity TA 35	Elbit Systems	292,002,936

[2] The analysis in this work used data till 2020.Q1.

3. Measures of Concentration and Similarity

Concentration Measures

We examine concentration segmented by Management company type, namely, Insurance companies, Pension funds, Mutual funds and Provident funds. We measure concentration in 3 different dimensions:

1. Concentration of Management companies
2. Concentration of Funds
3. Concentration of Assets

We aim to answer the following question: What is the effective number of companies, funds, and assets? The effective number differs from simple counting in the sense that effective number accounts also for size. To this end, we use the well-known indices: The Herfindhal- Hirschman Index (HHI) (Herfindhal, 1959; Hirschman 1945) and Entropy. These measures are calculated for each quarter and therefore we can examine their dynamics as a quarterly time series.

Herfindhal-Hirschman (HHI)

Assume for example that the aggregate portfolio of the management companies is invested in k assets with asset allocation represented by the following vector:

$$v = (v_1, \dots, v_k)$$

Where $v_i > 0$ is the total holding in asset i. Let,

$$p = (p_1, \dots, p_k)$$

be the vector of proportions of asset holdings where $p_i = \frac{\mathrm{v_i}}{\sum_{\{j=1\}}^{k} v_i}$. The HHI is defined by the squared norm of these proportions:

$$HHI = \left|\left|p\right|\right|^2 = \sum_{\{j=1\}}^{k} p_i^2$$

Entropy

The second measure we use is based on Entropy and is defined as follows:

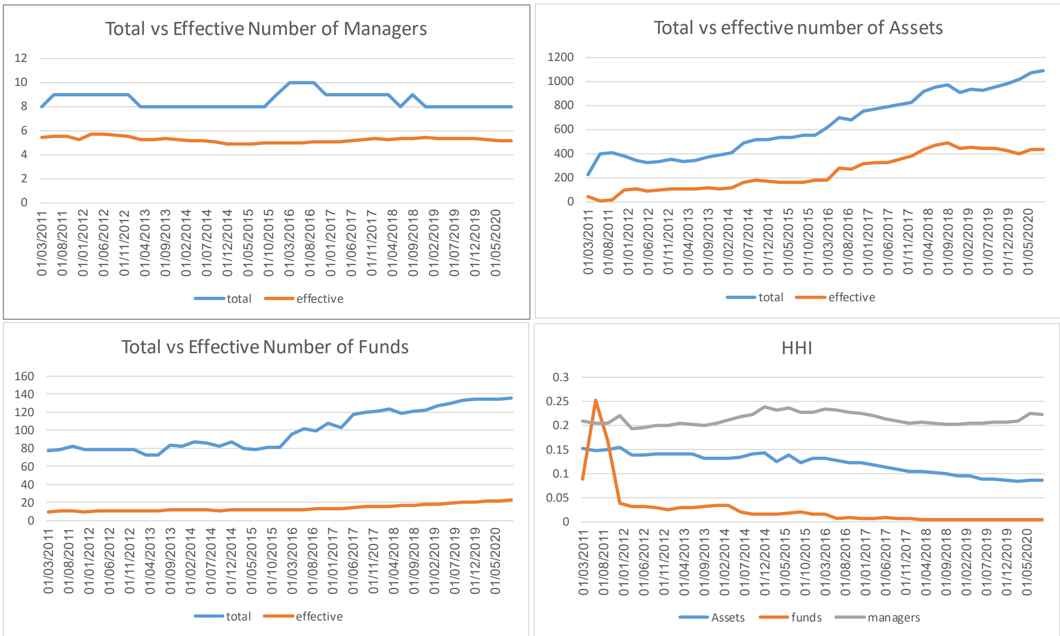
$$EFF = 2^{H(p)}$$

Where,

$$H(p) = -\sum_{\{j=1\}}^{k} p_i log_2(p_i)$$

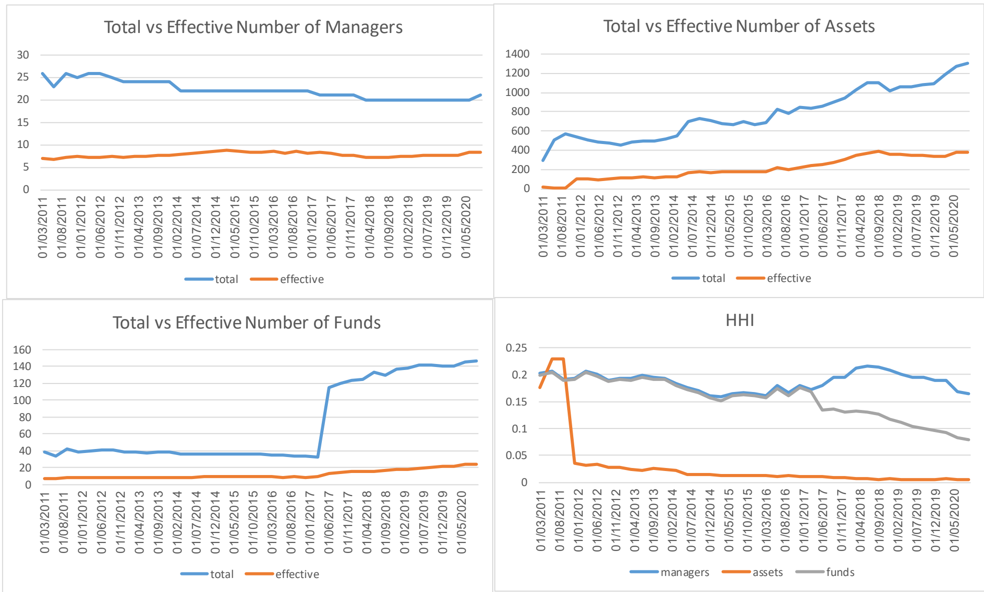
The Entropy is a common measure of diversity that gets values between 0 which points at full concentration and between $\log_2(k)$ which reflects full diversity in the case where $p_1 = p_2 = \cdots p_k = 1/k$.
These 2 measures are calculated in the same manner for companies, assets, funds and branches. For instance, when examining concentration of management companies k will be the number of management companies and $v_1, \ldots, v_k$ will be the vector of total assets that each company holds. In figures 1-4 we display concentration measures calculated in Insurance Companies.

Figure1 - Insurance Companies, Total vs Effective Numbers



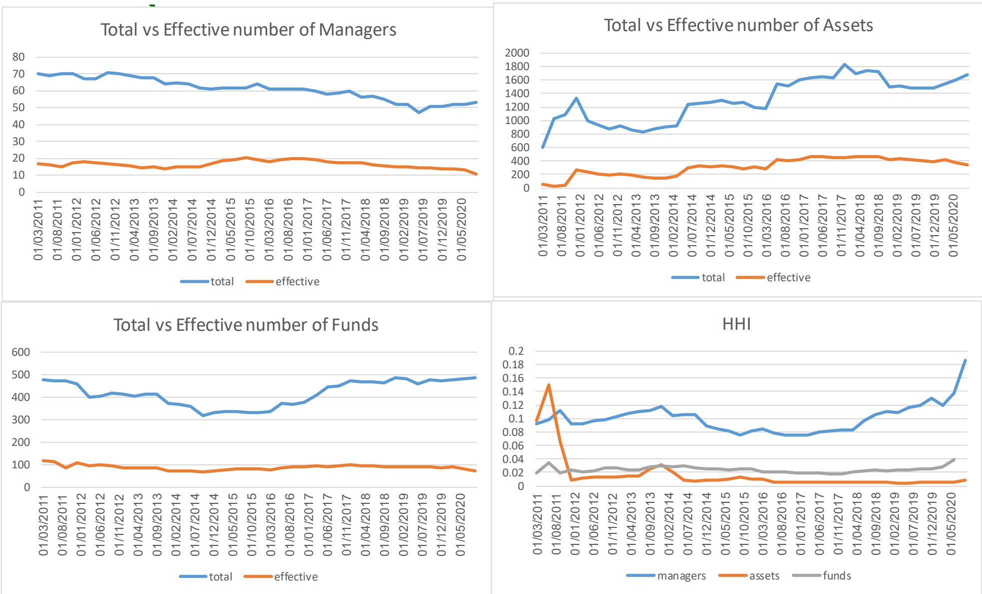
From figure 1 we can observe that the effective number of managers, assets and funds is significantly smaller (and smoother) than the total number. This indicates of higher concentration i.e. high percent of capital is concentrated in fewer managers, assets and funds, than the ones indicated by the total number of Managers, assets and funds, respectively. We can also observe a slight improvement across the years in diversity of the portfolio as indicated in the effective number of assets whereas there is no change in the effective number of managers. Moreover, there is a small effective number of Management companies which indicates of high concentration in the insurance industry.

Figure 2 – Pension Funds, Total vs Effective Numbers



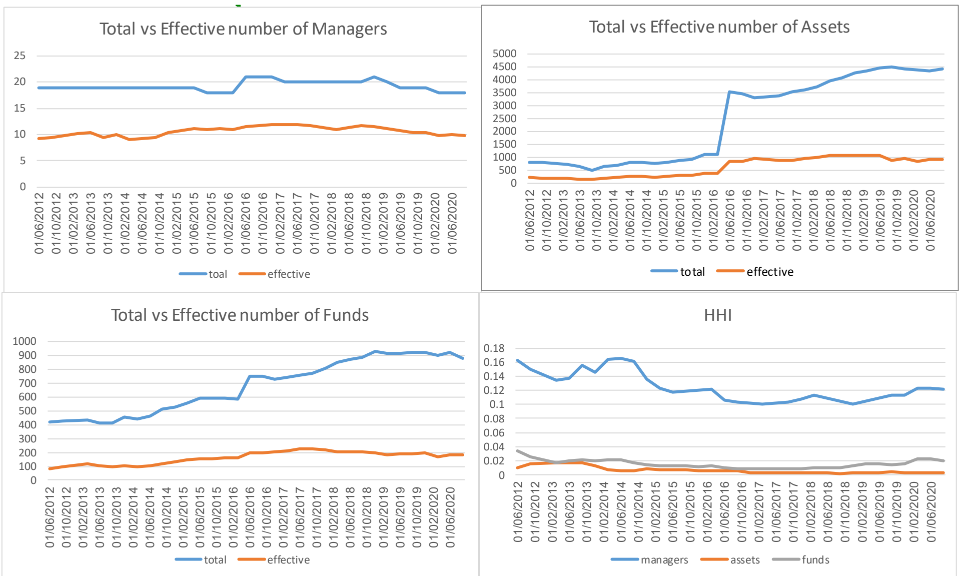
In the pension funds as can be observed in figure 2 there is also a slight improvement in the diversification of the asset portfolio. Also, in the pension field there is a small effective number of companies indicating high concentration. A structural break in 2017 is apparent in the effective number of funds. This is due to the change in reporting directives which required to report on specific investment channels rather than general management companies. This is apparent in the HHI index which shows the divergence between managers and funds beginning from 2017.

Figure 3 - Provident Funds, Total vs Effective Numbers



From figure 3 we can also observe a small number of managers and although there is an indication of worsening in the total number of companies the effective number indicates of stability in the number of managers. Also, the improvement in the number of assets is slighter than the one indicated by the total of assets.

Figure 4 -Mutual Funds, Total vs Effective Numbers



Similarity Measures

We examine similarity between portfolios of different management companies using Cosine similarity. To this end, we constructed for each management company the total holdings in each asset aggregated over the different funds. We then calculated between every 2 companies the cosine similarity measure. Cosine similarity assumes values between -1, which reflects absolute lack of similarity and 1 which reflects full similarity. In each quarter we represent these measures in a matrix that in cell i,j is the cosine similarity between company j and company i. These matrices can be constructed as heat maps where similar companies are close in the matrix. Following is a more formal definition of cosine similarity.

Cosine Similarity

Assume 2 management companies with the following vectors of total asset holdings in assets 1,…,k.

$$w = (w_1, \dots, w_k)$$
$$v = (v_1, \dots, v_k)$$

Where $w_i, u_i > 0$ the total asset holdings of each is company in asset i. The cosine similarity measure is defined between these two vectors is defined as:

$$Cos(v, w) = \frac{\sum_{\{j=1\}}^{k} w_i v_i}{\sqrt{\sum_{\{j=1\}}^{k} v_i}\ \sqrt{\sum_{\{j=1\}}^{k} w_i}}$$

Theoretically this measure can assume negative values but in our case, we only have non-negative asset holdings so this measures assumes only non-negative values.

Figure illustrates a heat map of portfolio similarity calculated with the cosine similarity measure for December 2014 in pension funds

Figure 5 – Heat Map of Cosine similarity between effective managers of Insurance Companies, 2019.Q1

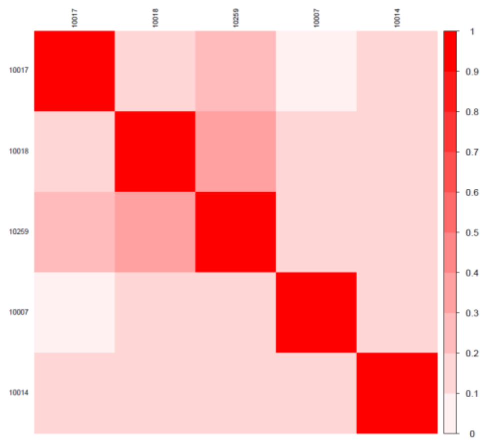


Figure 6 – Heat Map of Cosine similarity between effective managers of Provident funds, 2019.Q1

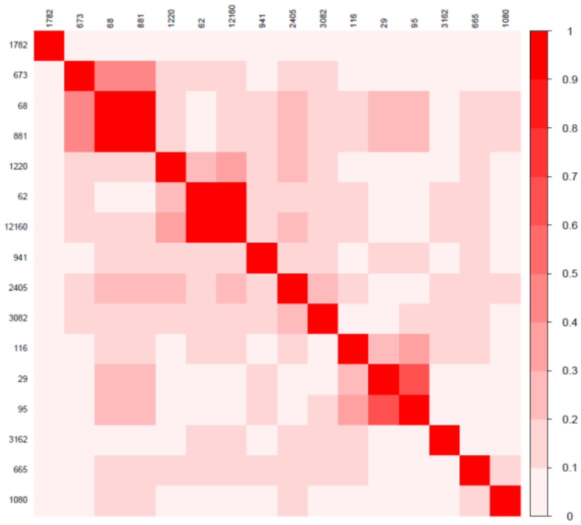


Figure 7- Heat Map of Cosine similarity between effective managers of Pension Funds, 2019.Q1

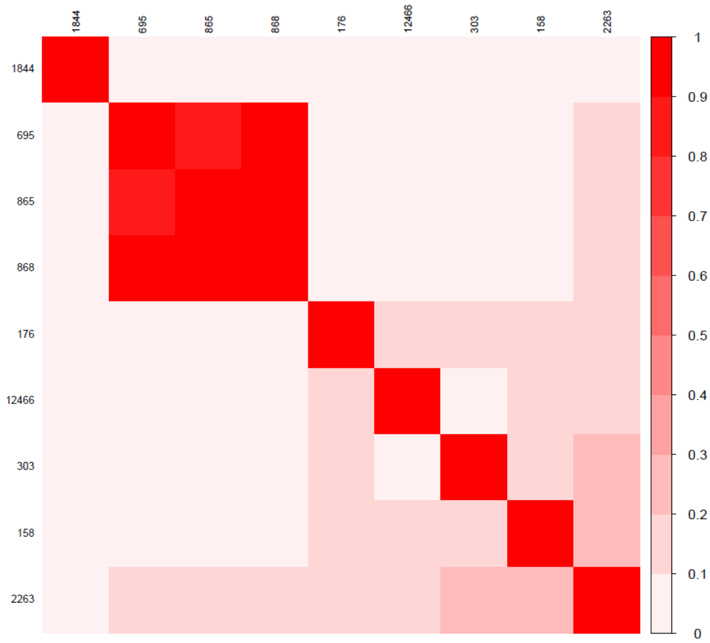
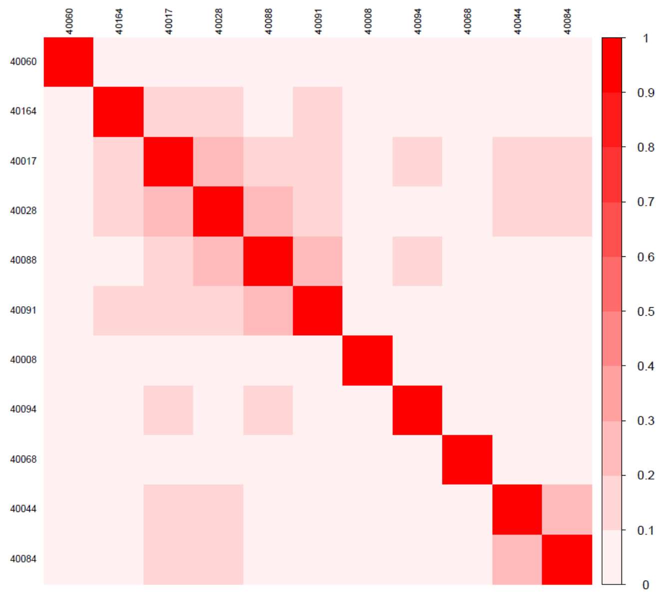


Figure 8 – Heat Map of Cosine similarity between effective managers of Mutual funds, 2019.Q1



The matrix of cosine similarity describes the portfolio similarity between each pair of companies.

4. Network analysis of institutional investor portfolios

A fully connected undirected weighted graph $G = (V, E)$ where $E = V^2$ can be induced from this matrix. In this graph each node represents a management company and edges between every pair of nodes is the cosine similarity between any two companies. We would like to filter this graph in a way that preserves the most relevant part of the data interdependency patterns measured by cosine similarity. We use for this purpose a method called Maximal Planar Graphs.

The Maximal planar graph is a filtering method of a dense graph or network. The main idea under these methods (see Massara et. al (2015)) is to filter a dense matrix of weights (correlations or cosine similarities) by retaining the largest and most significant sub graph while imposing constraints on the structure of the resulting graph. The minimal (or maximal) spanning tree retains the edges with the largest weights (correlations) while constraining the sub graph to be a spanning tree. The Maximal planar graph does the same while constraining the sub graph to be a planar graph (no edge crossing).

Figure 9 Network of Management Companies, 2014-12-31

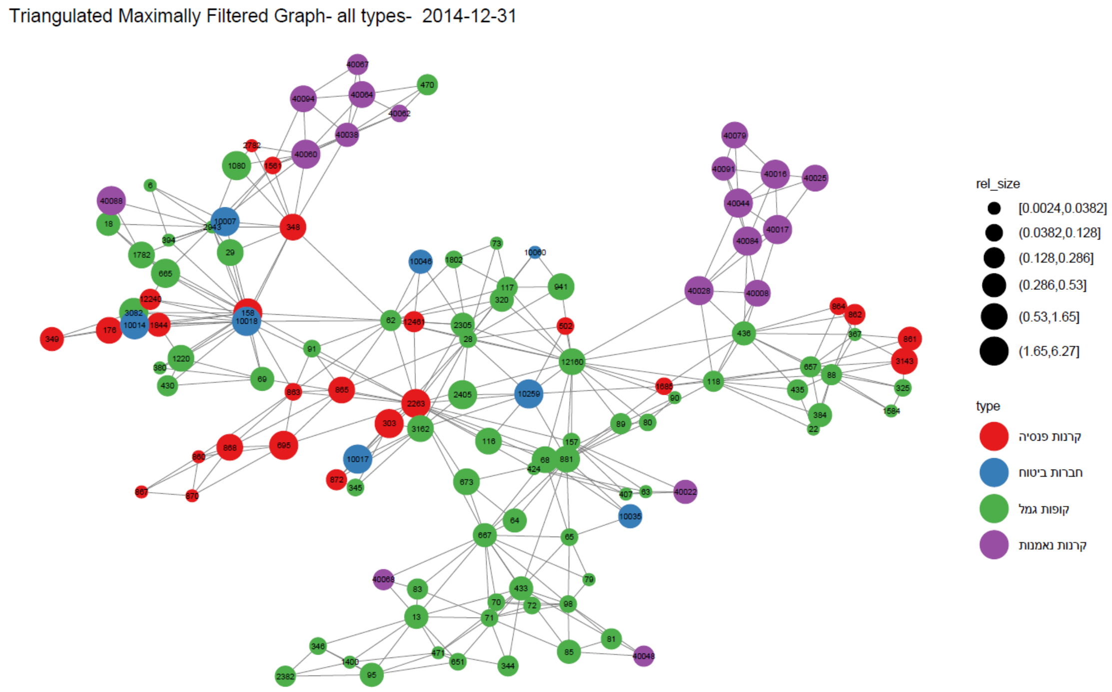
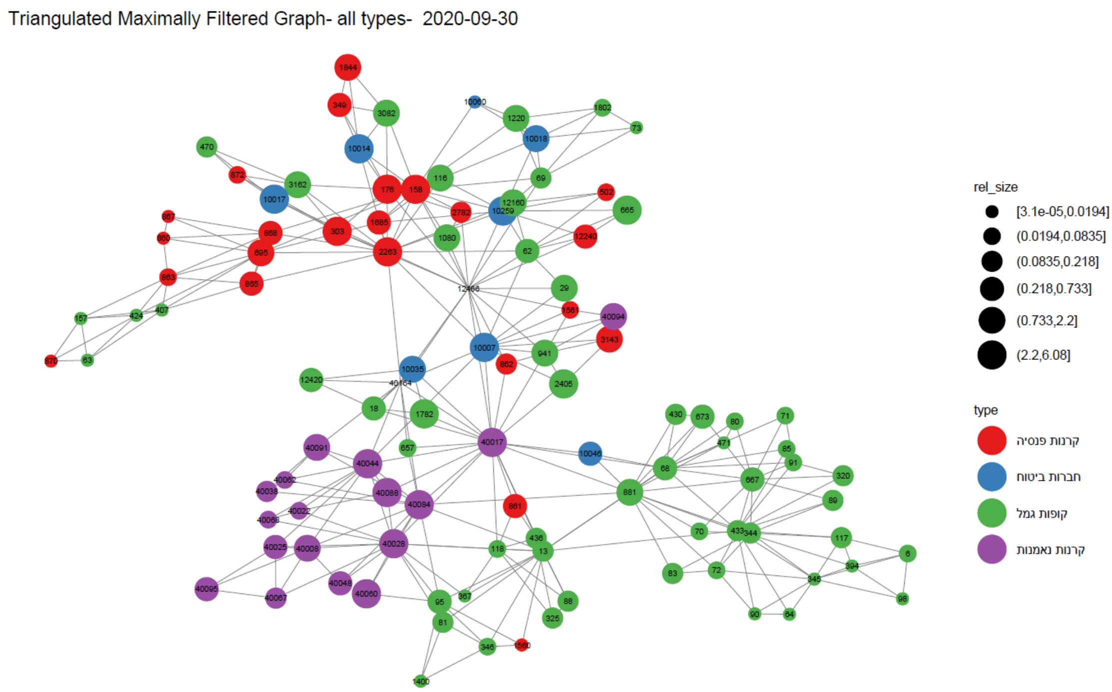


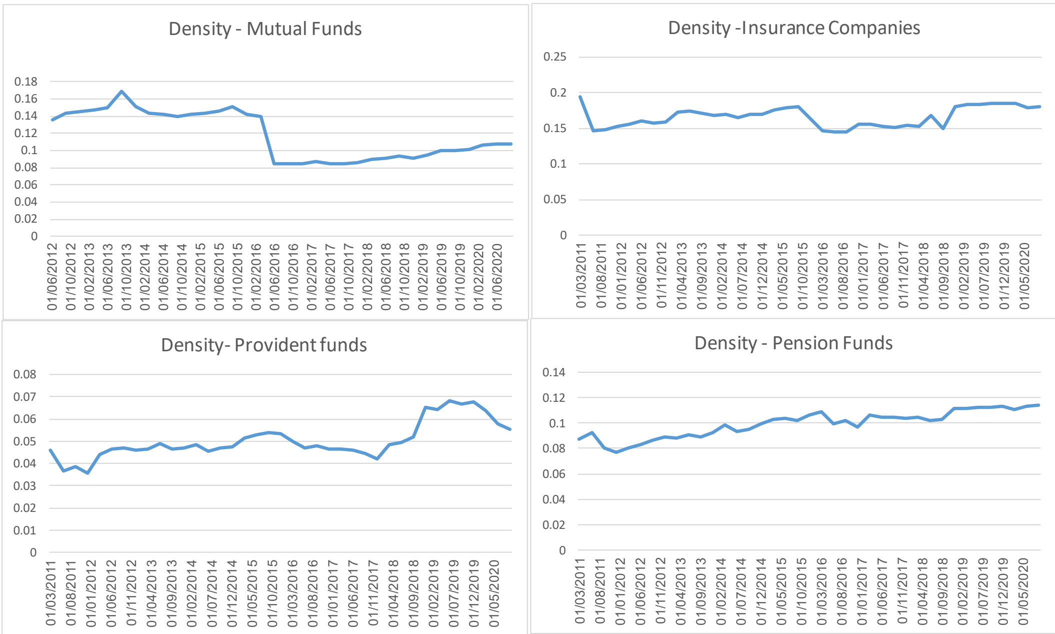
Figure 10 Network of Management Companies, 2020-19-30



We can observe from figure 10 that there is high similarity between management companies of the same type. The network exhibits clusters among provident funds and Mutual funds and also among pension funds. Insurance companies do not seem to cluster.

The density of a network is a measure of a graph is the ratio between the number of edges present in the graph and the maximum number of edges that the graph can contain. The density may reflect high aggregate similarity between management companies.

Figure 11- Density graphs by Management Company type



We can observe from figure 11 the structural break in the density around 2016 due to the inclusion of ETF's in addition to ETN's and hence weakening the density since 2016. There is an increase in the density among provident funds since 2017 due to the change in reporting directives as mentioned earlier.

References

Barucca, Paolo, Tahir Mahmood, and Laura Silvestri. "Common asset holdings and systemic vulnerability across multiple types of financial institution." Journal of Financial Stability 52 (2021)

Caccioli, Fabio, J. Doyne Farmer, Nick Foti, and Daniel Rockmore. "Overlapping portfolios, contagion, and financial stability." Journal of Economic Dynamics and Control 51 (2015): 50-63.

Elsinger, Helmut, Alfred Lehar, and Martin Summer. "Risk assessment for banking systems." Management science 52, no. 9 (2006): 1301-1314.

Girardi, G., Hanley, K.W., Nikolova, S., Pelizzon, L., and Getmansky Sherman, M. (2018). Portfolio similarity and asset liquidation in the insurance industry. SAFE WorkingPaperNo.224.

Herfindahl, O., 1950. Concentration in the U.S. Steel Industry, unpublished doctoral dissertation (Columbia University, New York, New York, U.S.A.).

Hirschman, A., 1945. National Power and the Structure of Foreign Trade (University of California Trade, Berkeley and Los Angles, California, U.S.A.).

Huang, Xuqing, Irena Vodenska, Shlomo Havlin, and H. Eugene Stanley. "Cascading failures in bi-partite graphs: model for systemic risk propagation." Scientific reports 3, no. 1 (2013): 1-9.

Levy-Carciente, Sary, Dror Y. Kenett, Adam Avakian, H. Eugene Stanley, and Shlomo Havlin. "Dynamical macroprudential stress testing using network theory." Journal of Banking & Finance 59 (2015): 164-181.

Previde Massara, Guido & Di Matteo, Tiziana & Aste, Tomaso. (2015). Network Filtering for Big Data: Triangulated Maximally Filtered Graph. Journal of Complex Networks. 5. 10.1093/comnet/cnw015.

Quantification of systemic risk from overlapping portfolios in the financial system
S Poledna, S Martínez-Jaramillo, F Caccioli, S Thurner Journal of Financial Stability 52, 100808

Upper, C. and Worms, A. (2004) Estimating Bilateral Exposures in the German Interbank Market: Is There a Danger of Contagion? European Economic Review, 48, 827-849.

Examining Concentration and Similarity in Institutional Investor's Holdings

Ariel Mantzura, Michael Bell
Bank of Israel

Data

- We use in this work Praedicta’s granular Data base.

- The data base includes the total asset holdings of management companies (including institutional investors) from 2010Q1 up to the present.

- The management companies are one of four: Insurance companies, pension funds, provident funds and mutual funds.

- The data is quarterly and detailed up to a particular asset and investment channel (funds).

- For each company and fund the total holding in assets is given.

Concentration: Questions?

- How many Management companies?
- How many effectively?
- How similar are the investment portfolios of the companies?
- How many assets? How many effectively?

Concentration Measures

- Assume that the aggregate portfolio is invested in k assets with asset allocation represented by the following vector:

$$v = (v_1, \dots, v_k)$$

- Where $v_i > 0$ is the total holding in asset i.

- Let,

$$p = (p_1, \dots, p_k)$$

be the vector of proportions of asset holdings where $p_i = \frac{\mathrm{v_i}}{\sum_{\{j=1\}}^{k} v_i}$.

- The **HHI** is defined by the squared norm of these proportions:

$$HHI = ||p||^2 = \sum_{\{j=1\}}^{k} p_i^2$$

- **Entropy**

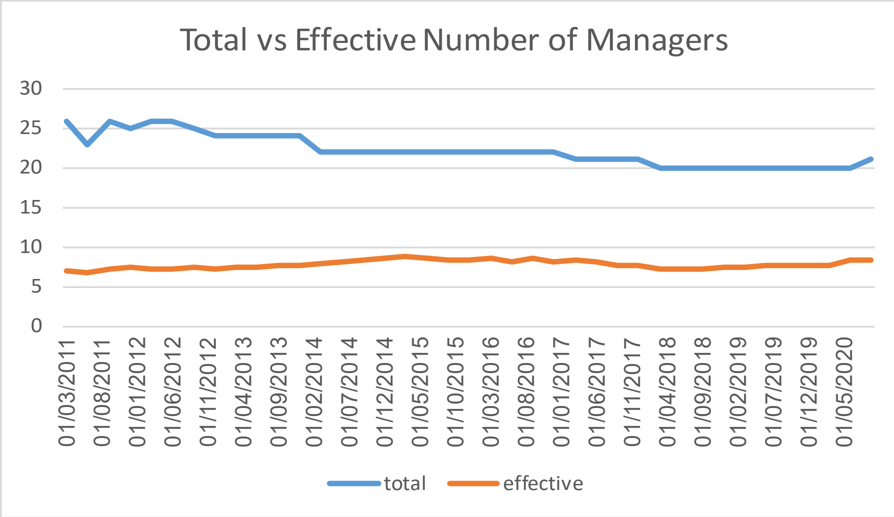
- The second measure we use is based on Entropy and is defined as follows:

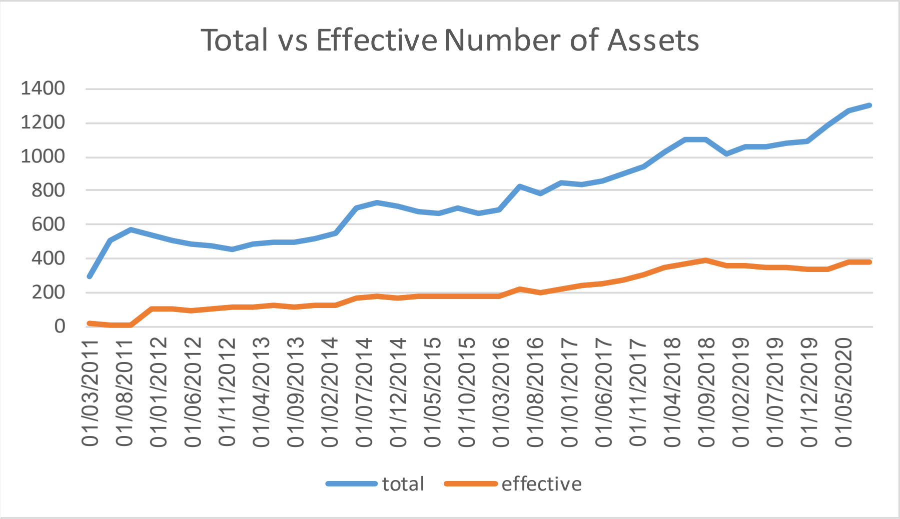
$$EFF = 2^{H(p)}$$

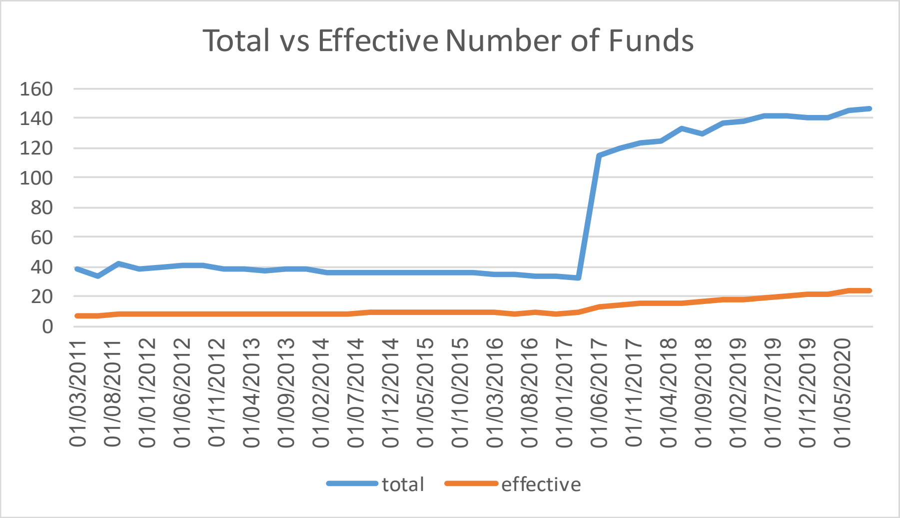
- Where,

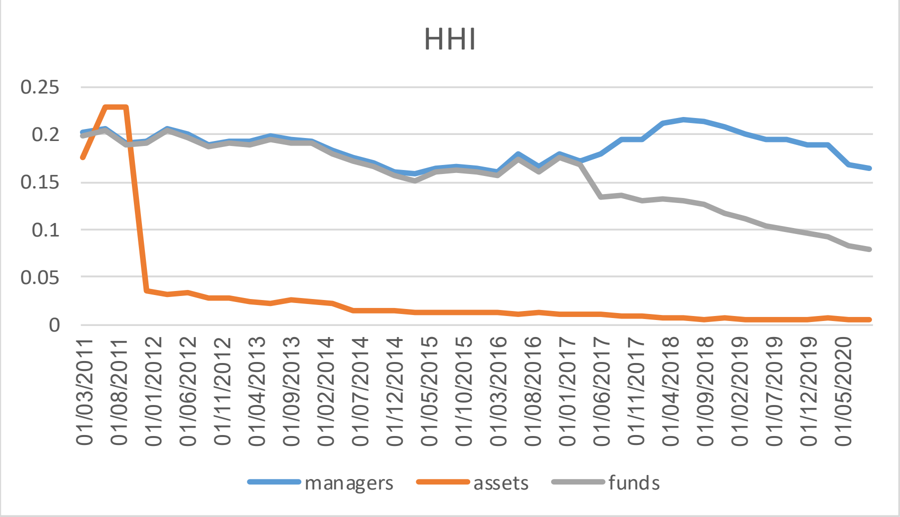
$$H(p) = -\sum_{\{j=1\}}^{k} p_i log_2(p_i)$$

Pension Funds, Total vs Effective Numbers









Portfolio Similarity – Cosine Similarity

- Assume 2 management companies with the following vectors of total asset holdings in assets 1,…,k.

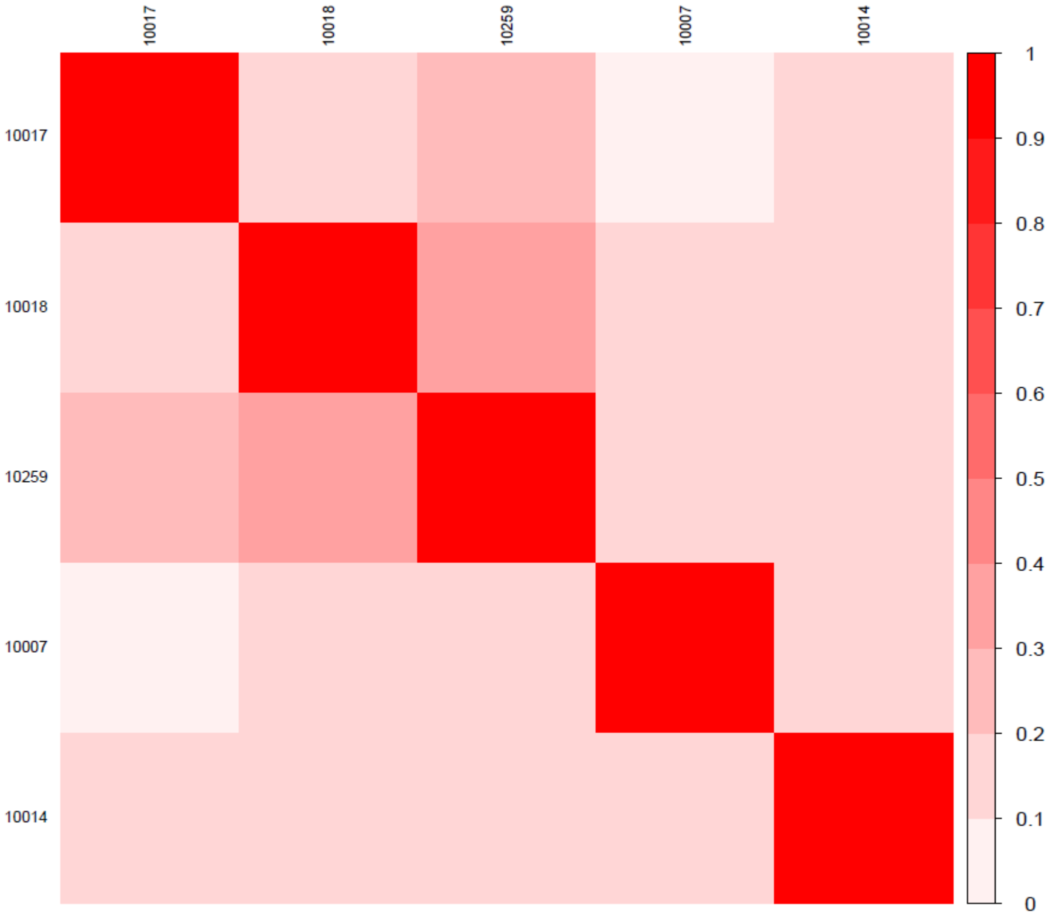
$$w = (w_1, \dots, w_k)$$
$$v = (v_1, \dots, v_k)$$

- Where $w_i, u_i > 0$ the total asset holdings of each is company in asset i.

- The cosine similarity measure is defined between these two vectors is defined as:

$$Cos(v, w) = \frac{\sum_{\{j=1\}}^{k} w_i v_i}{\sqrt{\sum_{\{j=1\}}^{k} v_i}\ \sqrt{\sum_{\{j=1\}}^{k} w_i}}$$

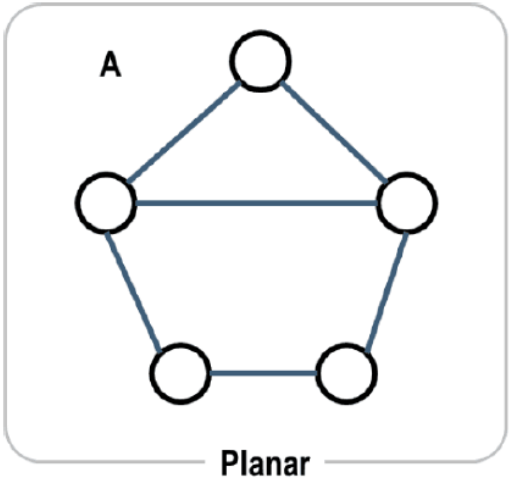
Heat Map of Cosine similarity between effective managers of Insurance Companies, 2019.Q1

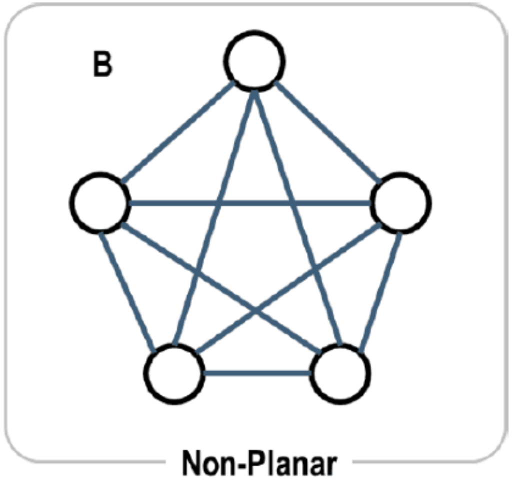


Filtering with maximal planar

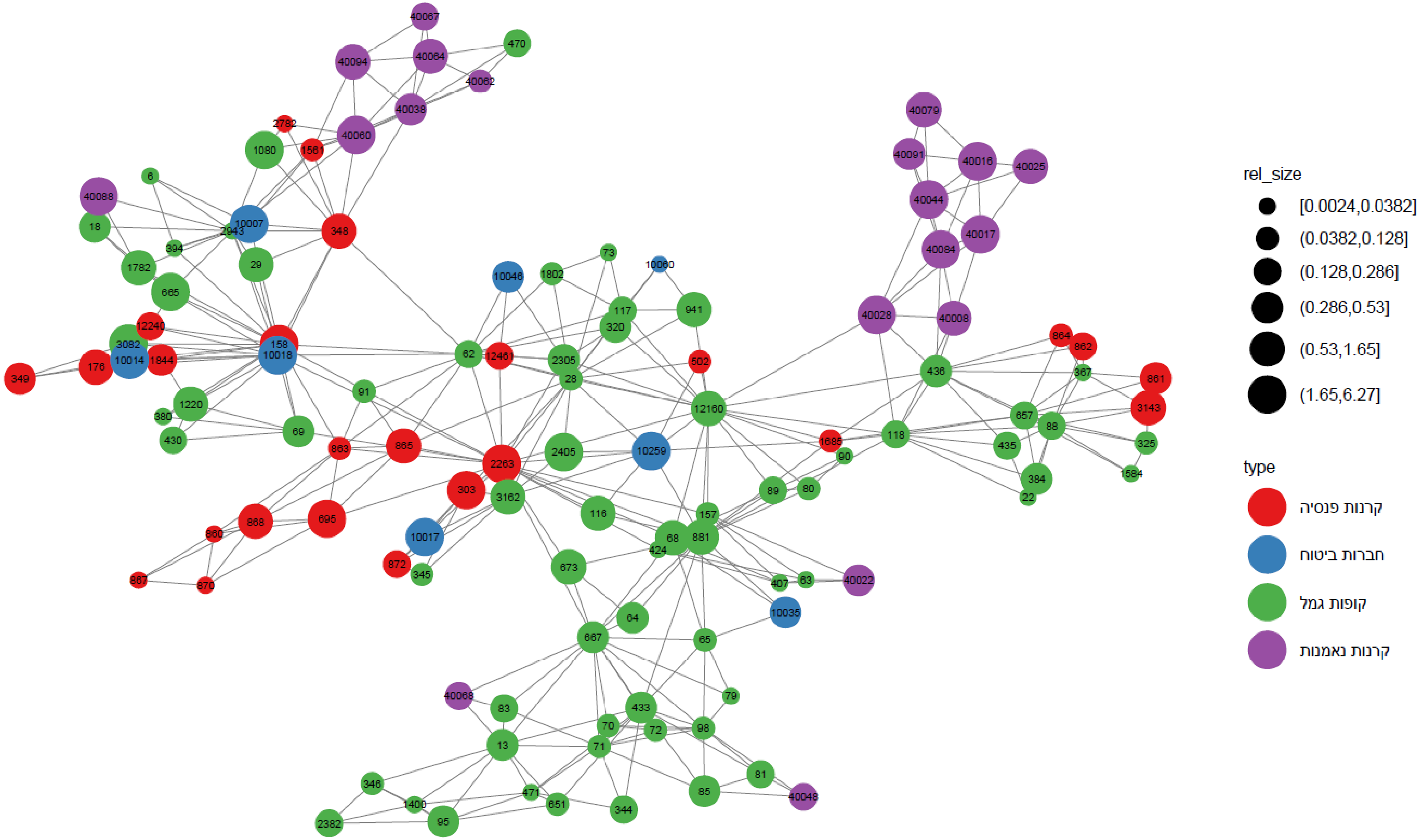
- The Maximal planar graph is a method of filtering dense matrix of weights (correlations or cosine similarities) by retaining the largest and most significant sub graph while imposing constraints on the structure of the resulting graph.

- The Maximal planar graph does the same while constraining the sub graph to be a planar graph (no edge crossing).





Triangulated Maximally Filtered Graph- all types- 2014-12-31



To Conclude

- The Effective Number is significantly smaller than the total number
- Significant portion of effective companies are highly similar
- Similarity is often present in same company type