

# 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  $p_i = \frac{v_i}{\sum_{j=1}^k v_j}$ .

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

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

- Entropy**

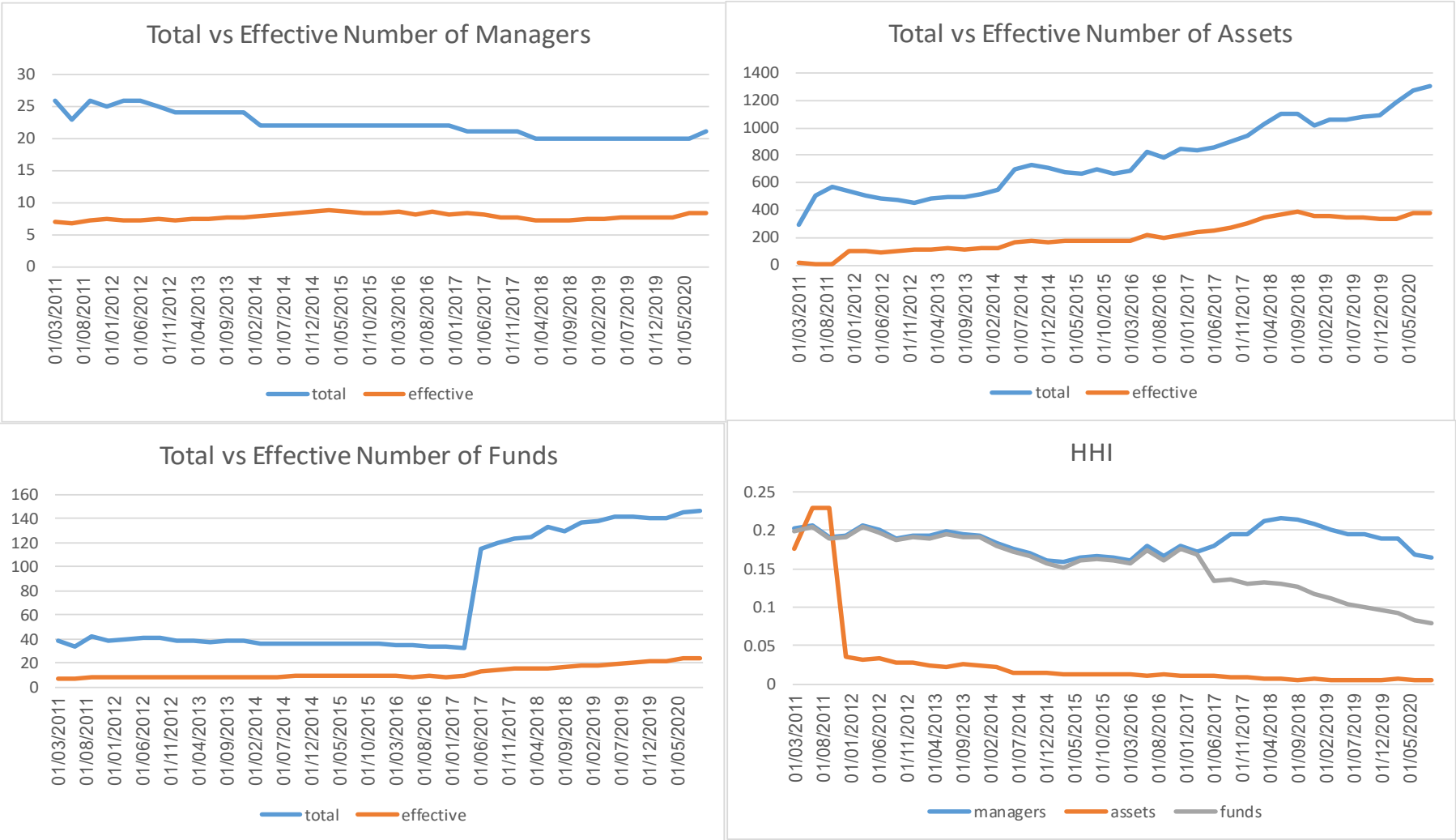
- The second measure we use is based on Entropy and is defined as follows:

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

- Where,

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

# Pension Funds, Total vs Effective Numbers



# Portfolio Similarity – Cosine Similarity



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

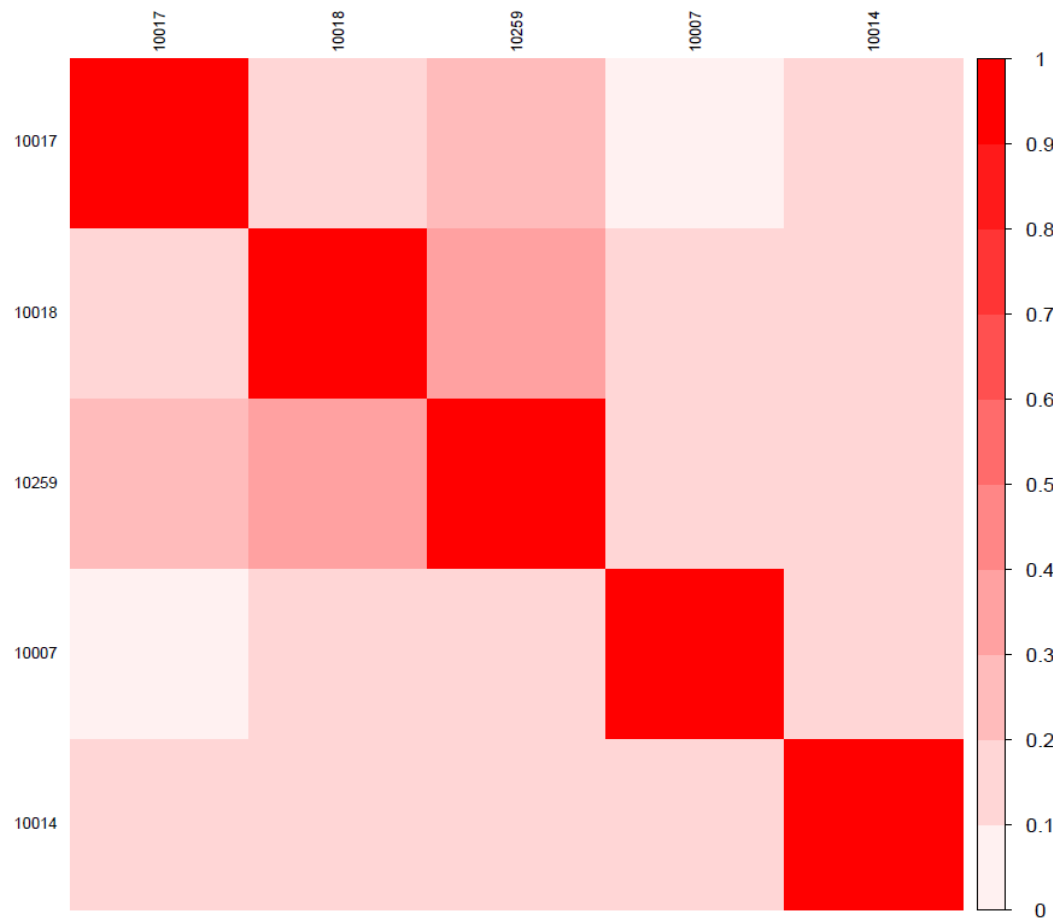
$$w = (w_1, \dots, w_k)$$

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

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

$$\text{Cos}(v, w) = \frac{\sum_{i=1}^k w_i v_i}{\sqrt{\sum_{i=1}^k v_i^2} \sqrt{\sum_{i=1}^k w_i^2}}$$

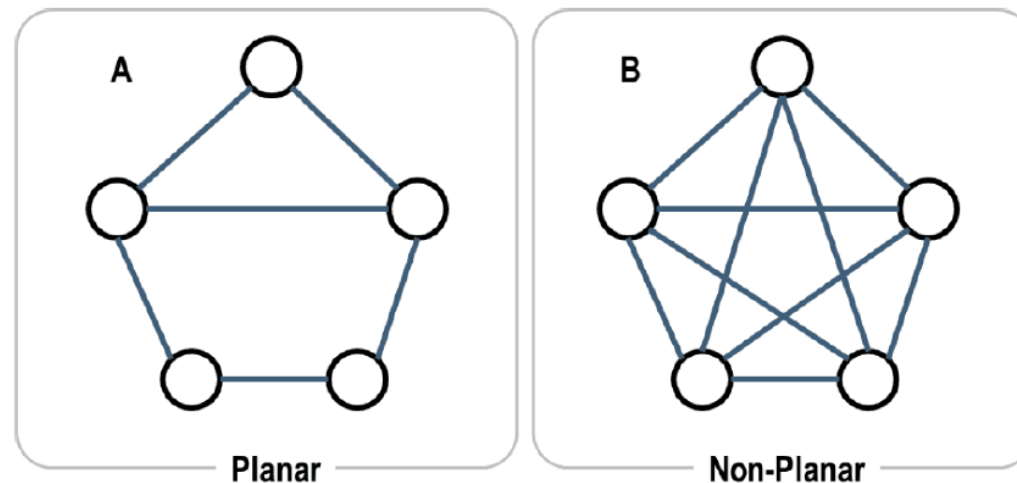
# Heat Map of Cosine similarity between effective managers of Insurance Companies, 2019.Q1



# Filtering with maximal planar

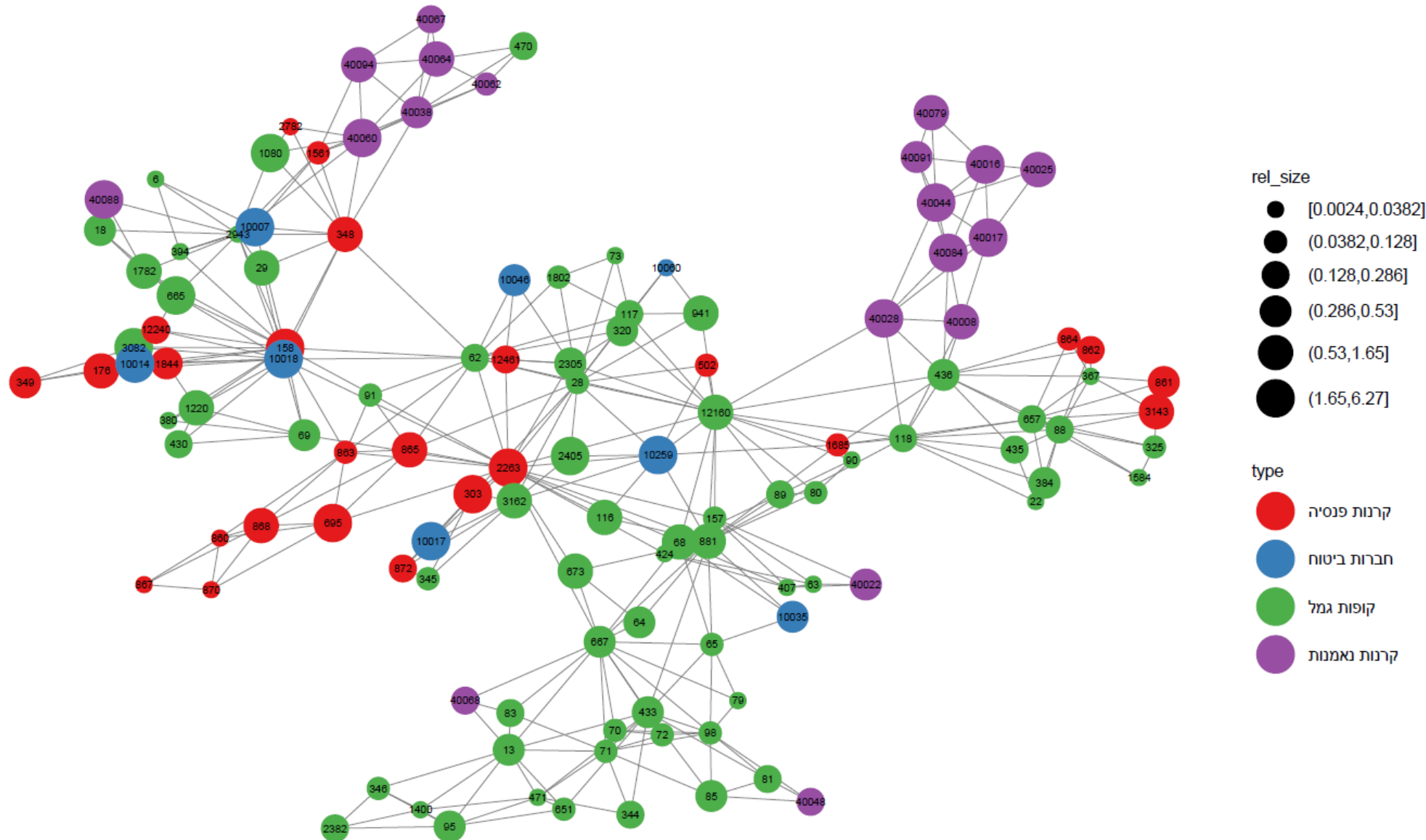


- The Maximal planar graph is a method of filtering dense matrix of weights (correlations or cosine similarities) by retaining the largest and most significant sub graph while imposing constraints on the structure of the resulting graph.
- The Maximal planar graph does the same while constraining the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