4

Economic Specialization and the Nested Bipartite Network of City–Firm Relations

Antonios Garas[1], Céline Rozenblat[2], and Frank Schweitzer[1]

[1]Chair of Systems Design, ETH Zurich, Weinbergstrasse 58, 8092 Zurich, Switzerland
[2]Geography Institute, Faculty of Geosciences, University of Lausanne, 1015 Lausanne, Switzerland

The set of goods and services produced in a city depend on a complex interplay of factors that include institutions, taxes, skilled personnel, industrial heritage, and the presence of particular resources. Dependent on the availability of such factors, some cities have specialized in certain economic activities, while others became economically more diversified. Specialization comes with a benefit as it allows for economic multiplier effects through "agglomeration economies" [215, 132, 160]. This, however, can turn into a drawback if a particular economic activity goes into recession. Then, cities specialized in this activity will be distressed more than the economically diversified ones. To exacerbate the problem, in a globalized world the economic performance of a city increasingly depends on the economic performance of other cities. Such economic dependencies emerge even between cities that are very far away in terms of geographical distance, which, due to global economic linkages, are, in reality, proximal in economic terms [256].

This leaves us with the problem of quantifying such dependencies and linking them to the diversification of economic activities in cities. In this chapter, we identify a city's economic activities by monitoring firms with global presence and which operate in a particular city. This way, we can create a network that links cities and firms and, if we focus on the economic activities of each firm, we can extend this network by linking cities to economic activities. Networks that describe realtions between different sets of nodes, in our case, cities and economic activities, are called bipartite networks (or bipartite graphs) [78]. The bipartite network we construct to link every city with the economic

Garas, A., Rozenblat, C., and Schweitzer, F., "Economic Specialization and the Nested Bipartite Network of City–Firm Relations" in *Multiplex and Multilevel Networks*, edited by Battiston, S., Caldarelli, G., and Garas, A.
DOI: 10.1093/oso/9780198809456.003.0004

activities of firms with global presence will allow us to study *how specialized or diversified* each city is with respect to a global context.

In general, firms are connected to other firms through different types of links, such as ownership relations, supply chains, and financial obligations. Likewise, cities are linked to other cities via transportation networks or financial networks. Therefore, we can represent all these relations using large, multilayered, interconnected network [107], where the different layers contain information about different kind of links connecting cities to cities, firms to firms, and cities to firms. Indeed, this is the most general way to describe any system with different interacting elements, using the complex networks framework. In this representation, the aforementioned bipartite network is just the part that describes the structure of the interlayer links that connect cities to firms, based on the firm's locations.

The structure of the city–firm bipartite network has striking similarities with other types of bipartite networks found in *ecology*. There, nodes represent species, while links their interactions. In so-called antagonistic networks, the interaction between species is asymmetric, such as host–parasite, predator–prey, and plant–herbivore interactions. In so-called *mutualistic* networks, on the other hand, the interaction between species is symmetric, that is, both species interact in a mutually *beneficial* way, such as, for example, the way that plants interact with their pollinators. Networks with antagonistic and mutualistic interactions have long been studied in ecology, to show that the stability of ecological communities is linked to structural features of the network topology [24, 293, 240]. More precisely, it was shown that mutualistic networks are organized in a nested pattern, while antagonistic networks are organized in compartments [293]. A nested organization means that the network consists of sets of generalist nodes and sets of specialist nodes. The specialists interact only with a small subset of nodes, while the generalists interact with (almost) all other nodes in the network. In nested ecosystems, the large set of interactions between generalists (i.e., species that interact with many other species) creates a dense core to which the specialists (i.e., species that interact with few other species) are attached. It was shown that the (empirically observed [21]) nested structure of mutualistic networks reduces the interspecies competition, which, as a consequence, allows ecosystems to support more species and increase biodiversity [24].

Thus, it would be of great interest to see whether such nested structures can be also found in bipartite networks related to economics. A recent study [129, 288] has investigated the bipartite network between firms and countries, to relate it to economic stability. It was found that robust countries have, indeed, a wide range of diversification in their economic activities. Based on this economic complexity, performance measures for countries were proposed. Specifically, it was shown that the dynamics of the nested structure of industrial ecosystems can predict path dependencies in the way industries appear and disappear in given countries. This helped to explain the evolution of the set of products that are produced and exported by these countries [248, 52]. A different analysis, focused on the New York garment industry, showed that a firm's survival probability depends on the firm's position in the nested network of interactions between designer and contractor firms [248].

In this chapter, however, we are interested in the bipartite network linking cities and the economic activities of globalized firms. These interactions are mutualistic because cities benefit from the operation of firms through taxes, employment, and so on, while firms benefit from cities through access to infrastructure, resources, customer bases, skilled personnel, and so on. To build our bipartite network of city–economic activity relations (an extract of which is shown in Figure 4.1), we used data about firms with global presence. For these firms, we know the precise locations of the headquarters and all their subsidiaries and, using the standard *Nomenclature of Economic Activities* (NACE), we can assign their core business to an economic sector (for details, see Section 4.1). Next, we studied the *structure* of this network, and we show that it follows a nested assembly, similar to ecological mutualistic networks. Therefore, building upon previous works in the field of ecology and their follow-ups with respect to economic networks, we show that ecological indicators can be used to identify the unbalanced deployment of economic activities, and we provide evidence that the structure of this bipartite network of city–firm relations contains information about the quality of life in cities.

4.1 The nested structure of links between cities and economic activities

For our analysis, we used data about the 3,000 largest firms with global presence and their $\sim$1 million direct and indirect links to $\sim$800,000 subsidiaries extracted from the BvD *orbis* database of 2010 [51]. The firm locations were aggregated using the concept of Functional Urban Areas (FUAs), which was developed by the *European Spatial Planning Organization Network* (ESPON) [87]. Using FUAs makes it possible to agglomerate municipalities according to their functional orientation—sometimes going beyond administrative boundaries—and reflect the actual operational conditions of people, enterprises, and communities. Therefore, FUA agglomerations result in an efficient mapping of the economic activity and service production. In addition, we classified firms into economic sectors according to their core businesses, using the NACE nomenclature provided by *eurostat* [89]. These sectors were further aggregated following the United Nation's *International Standard Industrial Classification of All Economic Activities* (ISIC REV.4) methodology, which results to an aggregation of activities into 21 different sections. Pairing the geographic location given by the FUAs with the NACE-ISIC classification of every individual firm, we created a bipartite network of interactions between 1,169 cities and 21 economic activities. An example of this network for the ten cities with the largest number of firms is shown in Figure 4.1.

Our bipartite network is represented by an incidence matrix M pairing each of the 21 economic activities to each of the 1,169 cities where this activity is present (see Figure 4.2). Therefore, each matrix element m_ij has a value of 1 if the economic activity i is present in the city j, and 0 otherwise.

To calculate the nestedness value of this matrix, we used the NODF algorithm developed by Almeida-Neto *et al.* [7]. This algorithm returns a nestedness value N in the range [0, 100], with $N = 0$ when there is no nestedness, and $N = 100$ for the case of perfect nestedness. To assess the significance of nestedness, we compared our measured value with a benchmark null model. In this chapter, our model of choice is the null model introduced by Bascompte *et al.* [21], which creates randomized networks by preserving

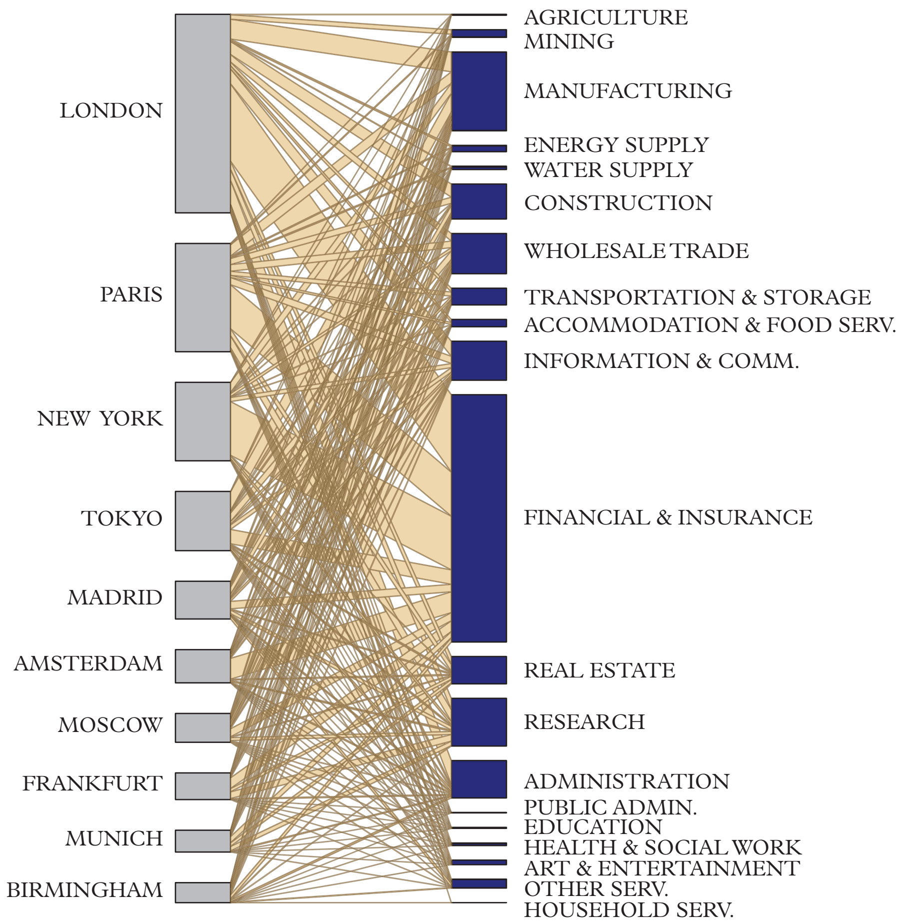

Figure 4.1 *City–economic activity mutualistic interactions. City–activity interactions as a bipartite network. The activity links of the ten cities with the largest number of firms are shown. The link width between city* i *and activity* j *corresponds to the number of firms associated to activity* j *located in city* i.

the degree distribution of the original network. This model generates ensembles of swapped matrices $\tilde{M}$ where the probability of each matrix cell being occupied is the average of the probabilities of occupancy of its row and column. Practically, this means that the probability of drawing an interaction is proportional to the level of generalization (degree) of both the city and the economic activity, that is $p_ij = (k_i/n_j + k_j/n_i)/2$, where k_j is the degree of the city, and k_i is the degree of the activity in the bipartite network, while n_i and n_j are the number of available activities and cities, respectively.

To measure the contribution of each individual city to the nestedness value of the whole network, we follow the methodology of Saavedra *et al.* [248]. More precisely, we calculate $c_i = (N - \langle N_i^* \rangle)/\sigma_{N_i^*}$, where N is the observed nestedness of the whole network,

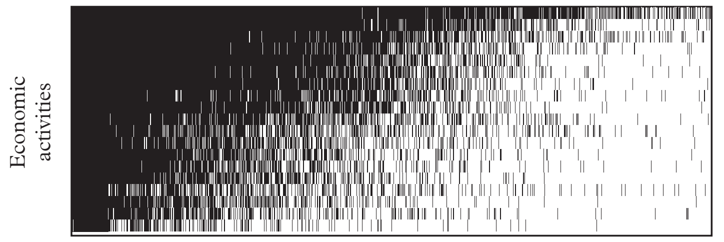

Figure 4.2 *The city–activity interaction matrix Plot of the interaction (incidence) matrix* M *pairing each of the 21 economic activities to each of the 1,169 cities where this activity is present. Each matrix element* m_{ij} *has the value 1 if the economic activity* i *is present in the city* j*, and 0 otherwise.*

and $\langle N_i^* \rangle$ and $\sigma_{N_i^*}$ are the average and standard deviation, respectively, of the nestedness across an ensemble of 100 random replicates for which all the links of city i to economic activities have been randomized. The number of random replicates is chosen in order to provide optimal performance, while at the same time the individual contribution to the nestedness of each city has converged significantly to their asymptotic value. More precisely, we performed a convergence analysis for which we calculated the Spearman's ρ correlation coefficient between two consecutive rankings of cities according to their nestedness contribution, with increasing numbers of replicates. From this analysis, we (a) observed that the rankings indeed converge to a saturation level and (b) concluded that 100 random replicates are enough, as ρ is already almost 0.99.

Using this methodology, we find that the bipartite network of cities–economic activities is nested (see Figure 4.2), with a nestedness value $N = 78.4$ ($p < 0.0001$). This already highlights structural similarities in the interaction patterns that occur in a natural ecological system and in the human-made economic system. And since a nested network structure is known to promote community stability in mutualistic ecological networks [293], we anticipate that the mutualistic network of cities and economic activities would be stable as well.

However, as was shown recently for both ecological and socioeconomic networks, nestedness comes at a price [248]. On the one hand, the nodes that contribute more to the nestedness of the network are the nodes that contribute more to the network persistence. On the other hand, these same nodes were identified as the ones most vulnerable to going extinct. Of course, in our case, a city might not go extinct, but it may decline to a less prosperous state.

4.2 Learning from ecology: Contribution to nestedness and economic well-being

As shown in Figure 4.3(a), the distribution of the individual contribution to nestedness is concentrated around the mean value $\mu = 1.96 \pm 0.01$. Therefore, one important question

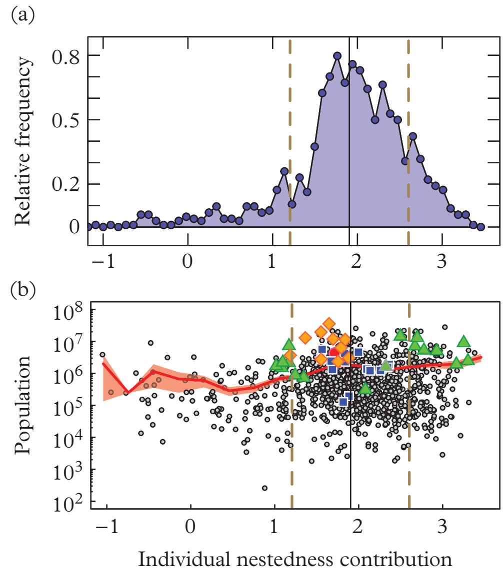

Figure 4.3 *Individual nestedness contribution. (a) Distribution of individual nestedness contribution for all cities. (b). Population versus individual nestedness contribution. The red line shows the mean value, and the band shows the standard error. The blue squares highlight the location of the top ten cities according to the Mercer 2012* Quality of Living *worldwide city rankings [191] and the top ten cities according to the EIU's 2013* Global Liveability Ranking and Report *[292] while the green triangles highlight the location (where available) of the bottom ten cities for both these rankings. In addition, the orange diamonds indicate the locations of the ten cities with the largest numbers of firms.*

is whether this nestedness value has any relation to a city's economic performance. If it does, then where are the best performing cities located in this distribution? Are they close to the center or close to its tail? Unfortunately, we do not have access to data about economic performance for individual cities. However, we do expect economic performance to be strongly correlated with the well-being of a city's inhabitants. In this sense, using the rankings provided by the Economist Intelligence Unit's (EIU) 2013 *Global Liveability Ranking and Report* [292] and the Mercer 2012 *Quality of Living* worldwide city rankings [191], we calculated the nestedness contribution of the top ten and the bottom ten cities. We found that, in both ranking systems, all of the top cities are within the range of $\mu \pm \sigma$, while 70% of the bottom cities are outside this range. More precisely, 41% of them are above $\mu + \sigma$, and 39% below $\mu - \sigma$. From Mercer's bottom-ten list, the cities above $\mu + \sigma$ are Abidjan, Khartoum, Kinshasa, and Conakry, while, from EIU's list, Karachi, Algiers, and Douala are in this range. In the area below $\mu - \sigma$, from Mercer's bottom-ten list we find Tbilisi, Sanaa, and Baghdad and, from the EIU's list, Damascus and Tripoli. This discussion shows that a ranking based on the nestedness score gives insightful results, where the better performing cities, according to Mercer and the EIU [191, 292] are closer to the mean of the nestedness distribution, while the worst performing ones are further away (see Figure 4.3(b)).

Given the general tendency of cities to grow, it is natural to ask if there is any measurable impact of population to the nestedness score. It is already known that a city's population drives many diverse properties of cities [31]. Are smaller cities more stable or more vulnerable, according to the way stability/vulnerability is reflected through nestedness? To answer this question, we collected data about cities' populations, by consolidating information taken from the United Nations database on cities,[1] the Organisation for Economic Co-operation and Development's database on cities,[2] and the ESPON project [87]. As shown in Figure 4.3(b), there is no pronounced relationship between the (logarithm of) population and individual nestedness. The Pearson correlation coefficient $r = 0.069$ ($p = 0.017$) is small and not significant, and the same is true for the Spearman correlation $\rho = -0.0052$ ($p = 0.8588$). Of course, if a city performs well and increases its inhabitants' well-being, it may become the target of internal or external migration flows and eventually increase its population. However, its network position—as measured with respect to nestedness—does not seem to be influenced by the population.

We can, of course, anticipate that if a city is specialized in an economic activity, it will prosper as long as the activity fares well. If this activity is hit by turmoil, or just underperforms with respect to other activities, this may lead to the city's decline. To avoid this, diversification of activities is required; but how much diversification is enough? And even if a city has indeed diversified its activities, how does this diversification compare to that of other cities? It is expected that large cities are able to attract many firms that would populate multiple economic sectors of activity [234]. This means that large cities are by definition "generalists" in the bipartite graph, and this introduces a bias in our interpretation of the nestedness score. To be more specific, let us consider the case of Detroit, which has the nestedness score $c = 1.71$, which places it near the mean of the nestedness distribution. Based on this number alone, we would argue that Detroit performs well, and we would not have anticipated its bankruptcy on July 18, 2013. Therefore, it is not enough to ensure that multiple economic sectors are populated; it is also important to monitor how many firms populate each sector, which is equivalent to using information about the weights of the links in the bipartite network. If the distribution of firms in economic sectors is skewed, one or a few sectors will dominate. Thus, a major decline in the dominating sector will have a major impact on the city's economy, and this will indirectly affect all the other sectors as well.

In addition, coming back to our previous discussion on the multilayered interconnected structure of the overall city–firm network, this effect will be even more pronounced in the presence of links between firms from different sectors. Even though we do not have data about such "hidden" links, it is not hard to imagine that many service-related smaller firms (e.g., subcontractors or advance production services) provide

[1] http://data.un.org.
[2] http://stats.oecd.org/Index.aspx?DataSetCode=CITIES.

support and depend on the function of the large firms that belong to the dominant economic sector. Therefore, the decline of this sector will create a cascading effect that is very hard to properly evaluate in the absence of detailed dependency data.

In the example at hand, from the 4,455 total large firms that were active in Detroit, 2,299 belonged to the manufacturing sector. The second most populated sector was Financial & Insurance, with 642 firms. We expect that many of these firms have strong ties to the manufacturing companies and would be affected if something went wrong in the manufacturing sector. However, since we cannot document these ties, for simplicity, we will assume that all sectors are independent.

Hence, to detect when one sector is overrepresented in the overall economic activity, we calculated the fraction f_τ, that is, the number of firms in the largest sector over the number of firms in all other sectors. If $f_\tau \leq 1$, the city is well diversified across activities, while if $f_\tau > 1$, a particular sector dominates the economy, and the city might be at risk. For our dataset, and under the assumption of sectoral independence, $f_\tau = 1.066$ for the case of Detroit, which indicates the city's vulnerability.

If more refined data were available, we could improve the calculation by dividing the number of firms of all sectors that would be significantly affected by a decline in the largest sector by the number of firms that would not be affected by this decline. However, such calculation cannot be performed with the datasets currently available. We would like to contrast our measure f_τ with other existing indexes for diversity. A well-known index used in ecology to quantify the biodiversity of a habitat is Simpson's index [266]. It is also known as the Herfindahl–Hirschman index in economics, where it is used to measure market concentration [130, 127]. We calculated Simpson's index for all cities in our database and found that the resulting concentration ranking is strongly correlated (Spearman's $\rho = -0.976$ ($p \leq 0.0001$)), with the ranking based on f_τ. But most available indexes, including Simpson's index, do not allow for the easy identification of a threshold value that discriminates well-diversified cities from those that are not well diversified. This, however, can be achieved by the $f_\tau = 1$ value in our case.

There is a limitation when applying our f_τ index to extremely specialized cities, as it diverges in cases where (mostly due to data limitations) only one economic sector is present. These cases are also identified by Simpson's index as being extremely specialized cities and are assigned a zero value. It is, therefore, better practice to exclude such pathological cases from our analysis. For this reason, we restricted our calculation of f_τ to the 100 cities with the largest number of firms in our database.[3]

As shown in Figure 4.4, most of the cities have a $f_\tau < 1$, which is evidence of a balanced development. However, there are some cities with f_τ values not only > 1 but even greater than the value for Detroit. As it happens, most of these cities, which include New York, Amsterdam, Zurich, and so on, are large financial centers, and this fact highlights the fragility of an economic model that is largely dependent on

[3] We calculated the Simpson's index for the set of 100 cities with the largest number of firms in our database, and again we found that the resulting concentration ranking is strongly correlated (Spearman's $\rho = -0.963$ ($p \leq 0.0001$)) with the ranking based on f_τ.

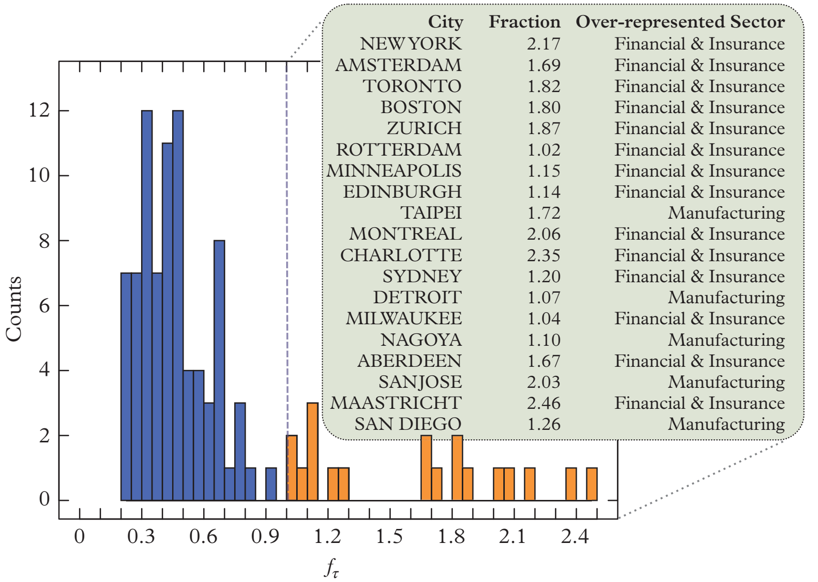

Figure 4.4 *Concentration of economic activities. Histogram of the fraction f_τ for the 100 cities with the largest number of firms in our database.*

financial services. The recent financial crises rang some bells, and now policy makers in developed countries are trying to mitigate this dependency by reshoring manufacturing. A profound example of this is the EU's 10|100|20 strategy, which aims to get almost 20% of semiconductor manufacturing back to Europe by 2020 through an unprecedented public–private investment partnership.[4]

4.3 Conclusion

In summary, by using the multilayered network approach and exploiting similarities in the bipartite network describing the organization of interlayer links across complex systems, we can use indicators developed in ecology to assess the performance of a city in the globalized economy. With these indicators, we go beyond the mere evaluation of the economic specialization of cities [134], as we associate the specialization of a city to the vulnerability of the whole "ecosystem" describing city–economic activities relations, similar to the way the extinction of one species affects the stability of natural ecosystems. Therefore, we highlight the possibility that such indicators have the potential to identify the need for new multilevel policies able to regulate the cities at the national or continental

[4] http://www.semi.org/eu/node/8506.

level (such as within EU), in order to enhance their position in the bipartite network of city–economic activities relations.

However, there is also a need for closer supervision to prevent the overrepresentation of some economic activities at the expense of others, as this increases risk in the future. In this respect, policy interventions that reduce the dominance of one sector over others should be applied more frequently. Currently, the financial sector is strongly overrepresented in most large cities in developed countries; hence, policies like the EU 10|100|20 strategy are important to hedge against future risk.

ACKNOWLEDGMENTS

The authors acknowledge financial support from the EU-FET project MULTIPLEX 317532. We thank Antoine Bellwald and Faraz Ahmed Zaidi for cleaning and aggregating the data. For our analysis, we used the R software for statistical analysis v3.0.2, and the bipartite library v2.02.