Appendix

A Copula definition and computation

As stated in the main document, the copula of random variables $(X_1, ..., X_d)$ corresponds to the joint cumulative distribution function (CDF) $C : [0,1]^d \to [0,1]$ of the uniformly distributed marginals $(U_1, ..., U_d)$. This means, that we first need to find the uniform multivariate distribution of the marginals. It is common practice to use the normalised ranked data, which corresponds to pseudo-observations defined as follows (see *pobs* function in R[1]):

"Given n realisations $x_i = (x_{i1}, ..., x_{id})^T$, $i \in \{i, ..., n\}$ of a random vector X, the pseudo-observations are defined via $u_{ij} = r_{ij}/(n+1)$ for $i \in \{1, ..., n\}$ and $j \in \{1, ..., d\}$, where r_{ij} denotes the rank of x_{ij} among all x_{kj}, $k \in \{1, ..., n\}$. This procedure ensures that the variates fall inside the open unit hypercube."

In order to make this research useful to the reader, we include below the algorithm used in this process for variables x_t and s:

```
library(VineCopula)
#The pseudo-observations are constructed as follows
set.seed(500)
u <- pobs(as.matrix(cbind(x_t,s)))

#Computation of the copula
selectedCopula <- BiCopSelect(u[,1],u[,2],familyset=NA)
```

The output of this is the family of the copula with the estimated parameters through maximum likelihood. For 2007 we obtained:

```
Bivariate copula: Survival BB1 (par = 0.13, par2 = 1.02,
                                                tau = 0.08)
```

where BB1 refers to the Clayton-Gumbel family, and survival is its rotation version by 180 degrees. We took the initiative to fit the data to the BB1 function instead of its survival counterpart, since this one is better documented. After implementation, we found that the correlation between the p_{ij} from the BB1 copula and its survival counterpart was bigger than 0.999. Hence the choice does not affect the results. BB1 functions are part of *Archimedean* copulas defined as follows (the following section is taken from [1]):

$$C(u_1, u_2) = \phi^{[-1]}(\phi(u_1) + \phi(u_2)) \tag{1}$$

where $\phi : [0,1] \to [0,\infty]$ is a continuous strictly decreasing convex function such that $\phi(1) = 0$ and $\phi^{[-1]}$ is the pseudo-inverse

$$\phi^{[-1]}(t) = \begin{cases} \phi^{-1}(t), & 0 \le t \le \phi(0), \\ 0, & \phi(0) \le t \le \infty, \end{cases}$$

[1] https://www.rdocumentation.org/packages/VineCopula/versions/2.2.0/topics/pobs

where ϕ is the *generator function* of the copula C.

The generator function for the BB1 is given by

$$\phi(t) = (t^{-\theta} - 1)^{\delta}, \qquad \theta > 0, \delta \geq 1. \tag{2}$$

The parameter τ obtained from the fitting (see result from code) can be deduced from θ and δ as follows: $\tau = 1 - 2/(\delta(\theta + 2))$. The BB1 copulas fitted for the 2007 and 2014 data gave rise to the following parameters: $\theta_{2007} = 0.028$, $\delta_{2007} = 1.069$, $\theta_{2014} = 0.013$, $\delta_{2014} = 1.061$.

B Percolation using copula for e^{-t} and s

In the following subsections we look at the effect of constructing the network for the set of nodes V given by the LSOAs, but using different weights for the links.

In this section we examine the effect of choosing a different function for time instead of $1/t$. Within transport modelling and spatial interaction models, the cost function is many times defined as either $1/t^{\beta}$ or as $e^{-\beta t}$. Let us therefore construct the copula using $x'_t = e^{-t}$ instead of $1/t$.

Repeating the same steps as explained in the methodology, we obtain as the best copula fit, the same copula as for the case where we used $x'_t = 1/t$, see Fig. 1.

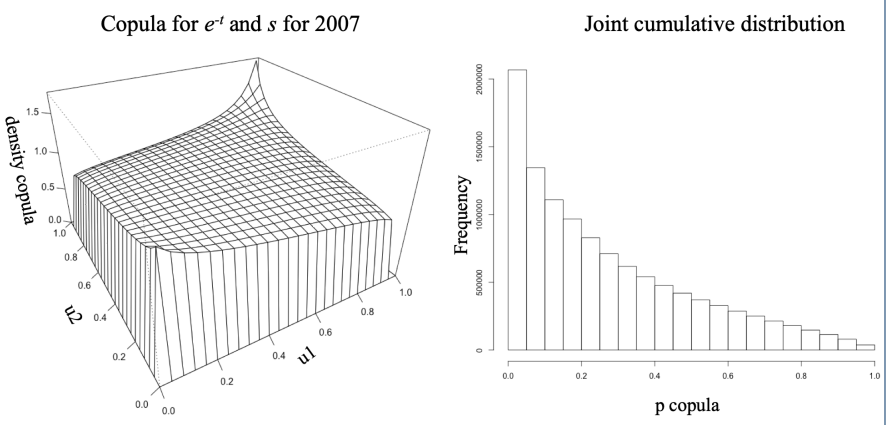

Figure 1. Copula for e^{-t} and s for London 2007. We see that this is the same as for $1/t$ and s in Fig.4 of the article

It should not be surprising that the results are the same for the copula constructed using the two functions $1/t$ and e^{-t}. The copula is modelling the dependencies between the industrial similarities of the LSOAs and how fast one can get from one to the other. Specifically, we are looking for pairs of LSOAs that have a good connectivity in terms of public transport, and that have a similar

ecosystem of industries. A shorter time t means a higher connectivity, and a higher s means high similarity, hence in order to have both variables reflecting a strong weight when both are high, etc, we need to transform t into a new variable $x'_t = f(t)$, so that a high x'_t corresponds to a strong connectivity. It is important to remark, that the two functions, $1/t$ and e^{-t}, keep the ranks of the transformed variables in a similar way (they keep the positive sign and the order of observations). Therefore, since the copula function uses the normalised ranks of the pairs of variable values t and s, rather than the raw values themselves (see how the pseudo-observations are generated in section A), the way time distances are transformed in our case (through $1/t$ or through e^{-t}) does not impact the results of the Copula and the clusters produced as a result.

C Percolation using only similarity

Let us now explore what would happen if we were to cluster the LSOAs according to their similarity only. In this case, we only consider the variable s for the percolation process.

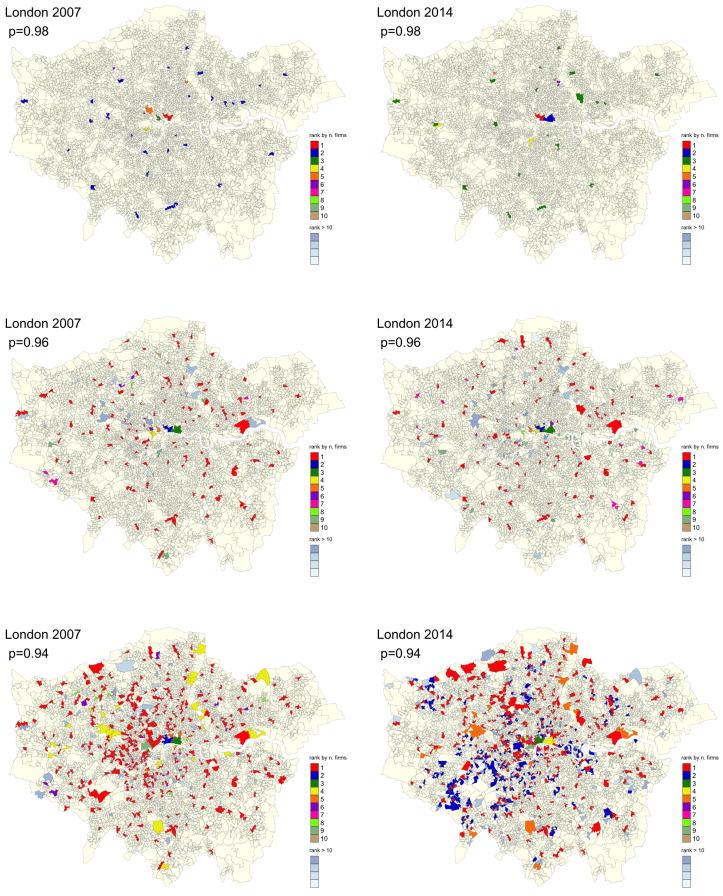

Figure 2. Geoindustrial clusters in 2007 and 2014 using only similarity s.

The Fig. 2 shows that different clusters are obtained for the different years, nevertheless, the clusters are composed of units that seem to be spread all over London, and hence do not look anything like *geographical clusters*.

D Percolation using only time

In this section, the links of the network are only defined by the different functions for the travel time. The travel time network was constructed using the timetables for May 2016, and we are using this network for 2007 and 2014, hence we only need to investigate the clusters for one year to illustrate how it works. Note though, that given that the cluster size is defined according to the number of firms inside each cluster, a possible variation in the colours might exist between the two years. This is however not relevant for comparison purposes with other methodologies.

Results for t, $1/t$ and e^{-t}

Running the percolation process on the network defined by the travel time only, t, $1/t$ or e^{-t}, gives rise to clusters that correspond to accessible zones at different time scales, see Figs. 3

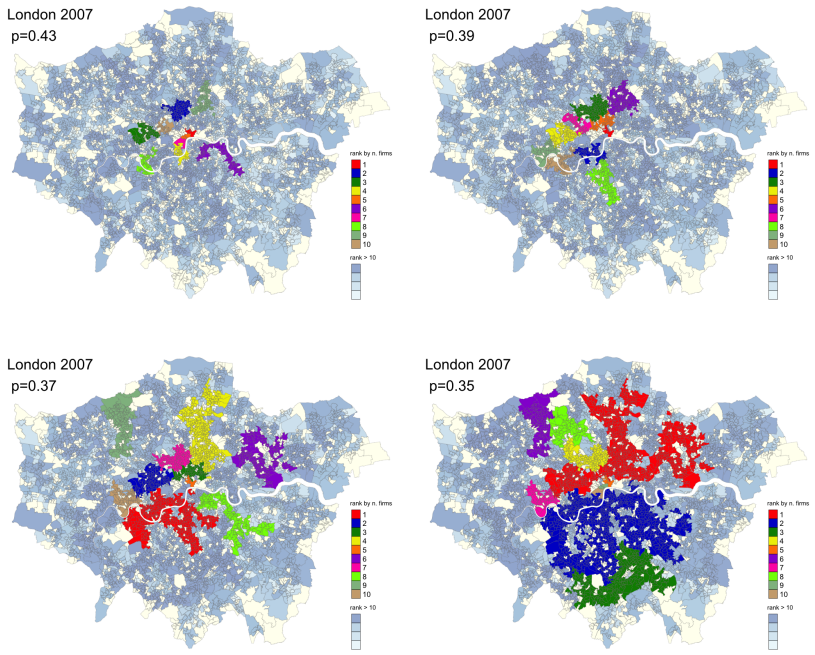

Figure 3. Geoindustrial clusters using only time $1/t$.

Overall, this exercise shows similarity or proximity given by time alone lead to results that are not satisfactory. The copula captures both aspects of the urban system. Of course further refinement can be obtained by using a more adequate measure of similarity or commuting flows.

E Diversity and concentration of firms

Diversity defined by Shannon's entropy at the level of clusters (Fig.4) recalls the general pattern of Fig.3 in the article (where SEI is computed at the level of LSOAs) whereby central London, at the exception of Temple, appears highly diverse in terms of industries represented by business units. The aggregation at cluster level smooths the variations in the West, but it highlights the presence of high diversity clusters like Croydon or the City of London: i.e. clusters which are made of close LSOA with a similar mix of diverse business units.

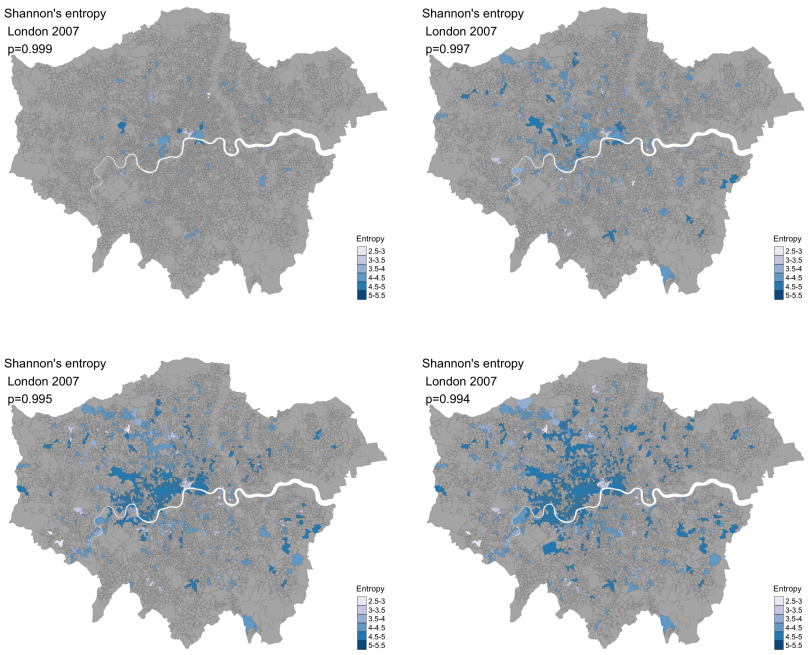

Figure 4. Shannon's entropy of the clusters defined at a specific threshold p for the copula of $1/t$ and s, London 2007.

Using the HHI index of specialisation at the level of clusters, we mostly single out the small and specific cluster of Temple Fig.5, very specialised in law activities. Other specialised clusters tend to be located in the periphery of London.

Author details

References

1. Brechmann, E. and Schepsmeier, U. (2013). Cdvine: Modeling dependence with c-and d-vine copulas in r. *Journal of Statistical Software*, 52(3):1–27.

Figure 5. Hirschman Herfindahl's index for the clusters defined at a specific threshold p for the copula of $1/t$ and s, London 2007.