Firm collaboration: exploratory analysis of FVG companies network

M. Franzon T. Rodani
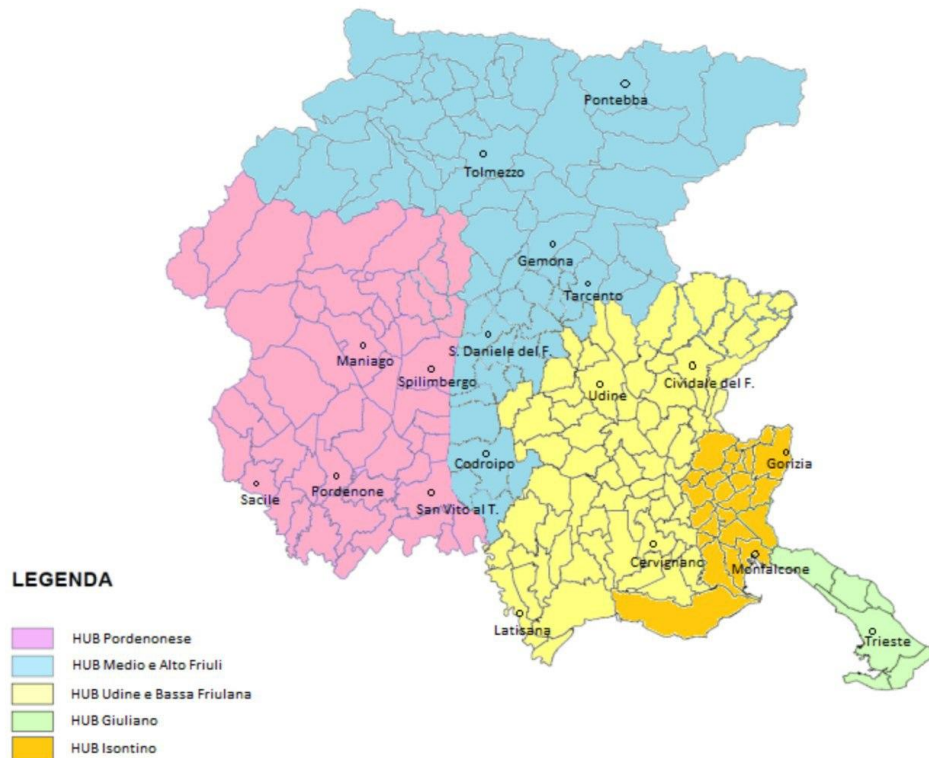
TABLE OF **CONTENTS**

01

DATASET

Firm collaborations between companies in FVG

DATASET

- **ATECO**, the national version of the EU NACE, a statistical classifications of economic activities.

- **SLL**, the national version of the EU Labour Market Areas (LMAs). **LMAs** are sub-regional geographical areas where the bulk of the labour force lives and works, and are usually different than administrative boundaries

- **Province,** which represent the institutional bodies of second level in the Italian Republic

- **HUB**, which are geographical areas that refers to different venues of "Agenzia regionale per il lavoro", the regional agency that offers labour services.

LEGENDA

- HUB Pordenonese
- HUB Medio e Alto Friuli
- HUB Udine e Bassa Friulana
- HUB Giuliano
- HUB Isontino

DATASET **ATECO**

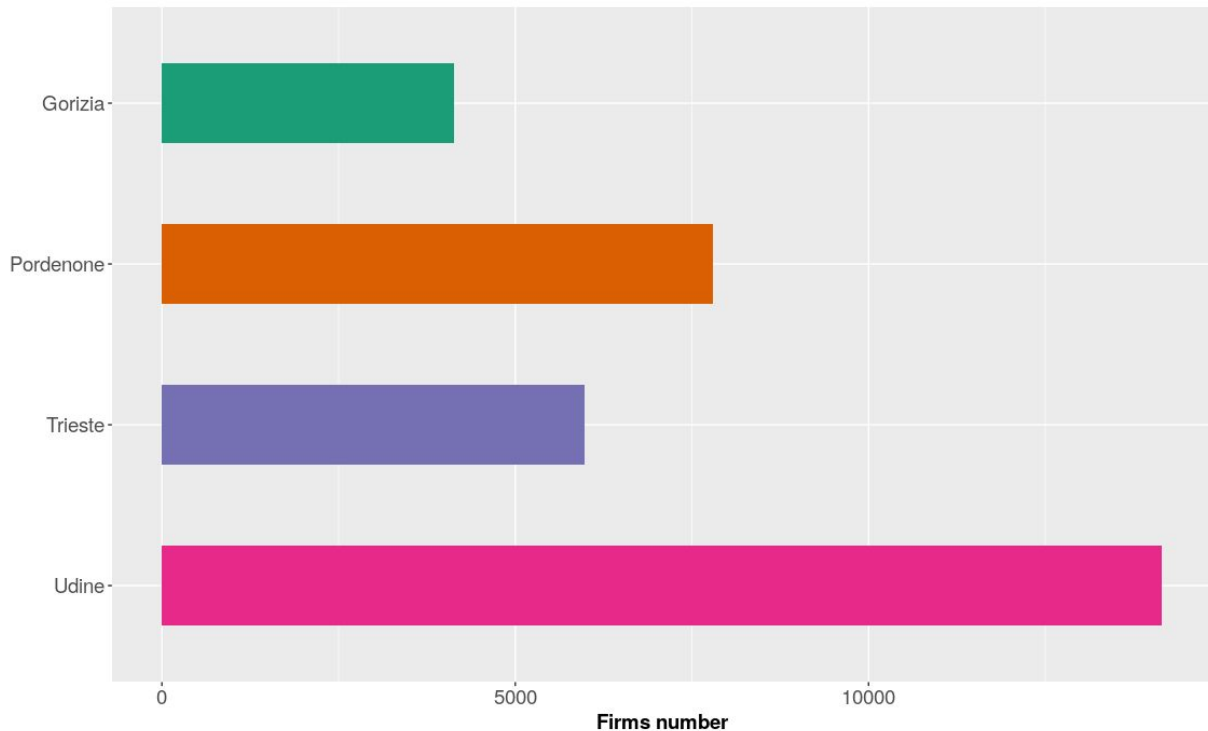**Unbalance** towards food services, retail and manufacturing

Firms distribution among ATECO codes in Friuli-Venezia Giulia
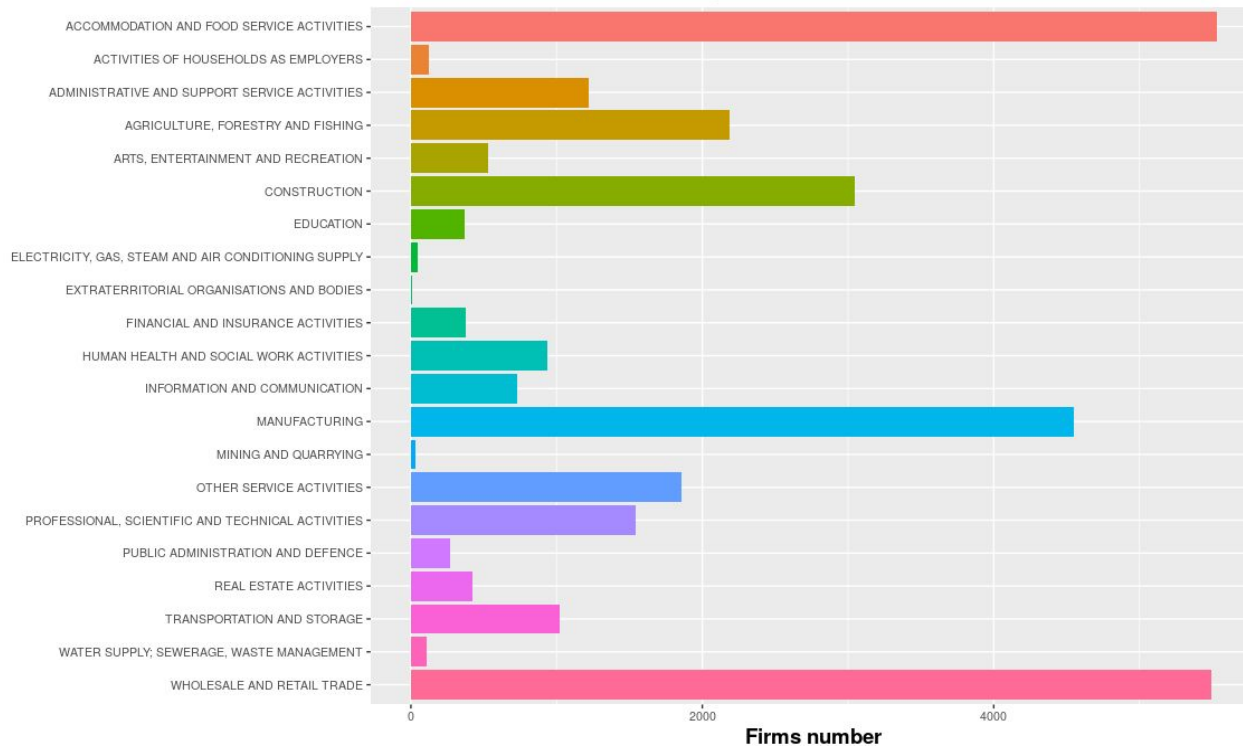
DATASET **PROVINCE**

Most of the firms are located in **Udine** province, which is the most populated
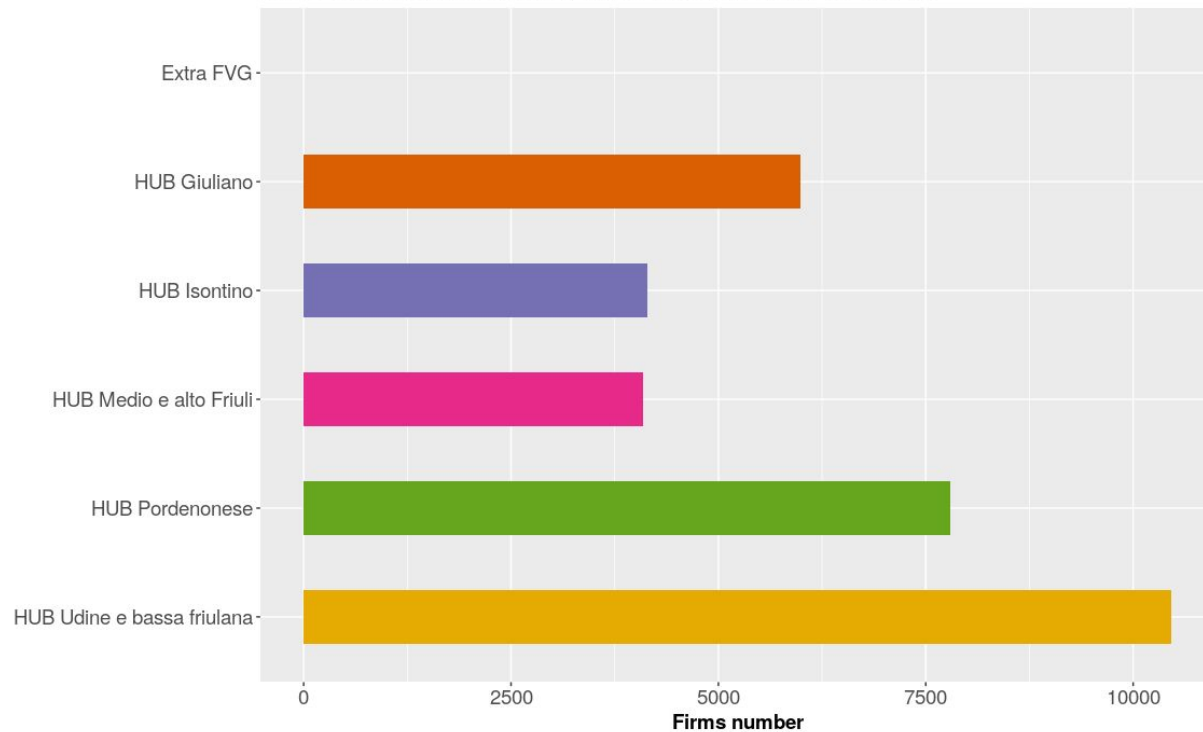
Firms distribution among province in Friuli-Venezia Giulia

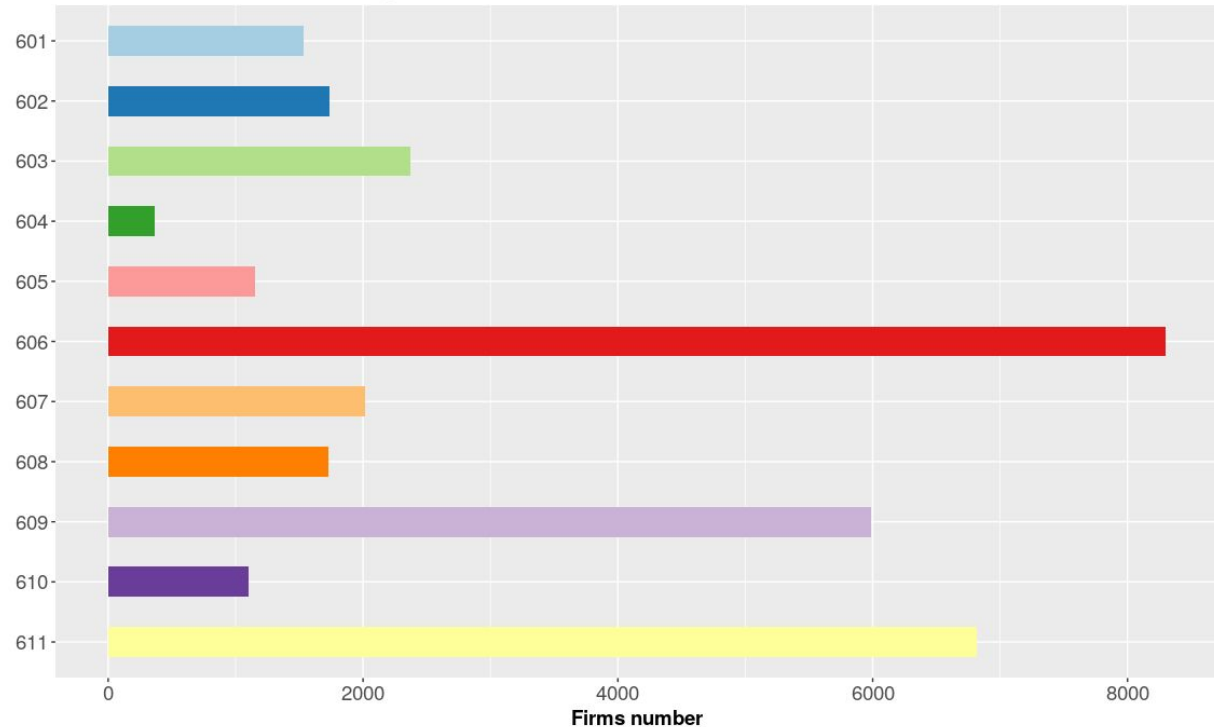# DATASET **HUB**

Extra FVG has only two firms

Firms distribution among hubs in Friuli-Venezia Giulia

DATASET **SLL**

606 stands for Udine; 611 for Pordenone and 609 for Trieste
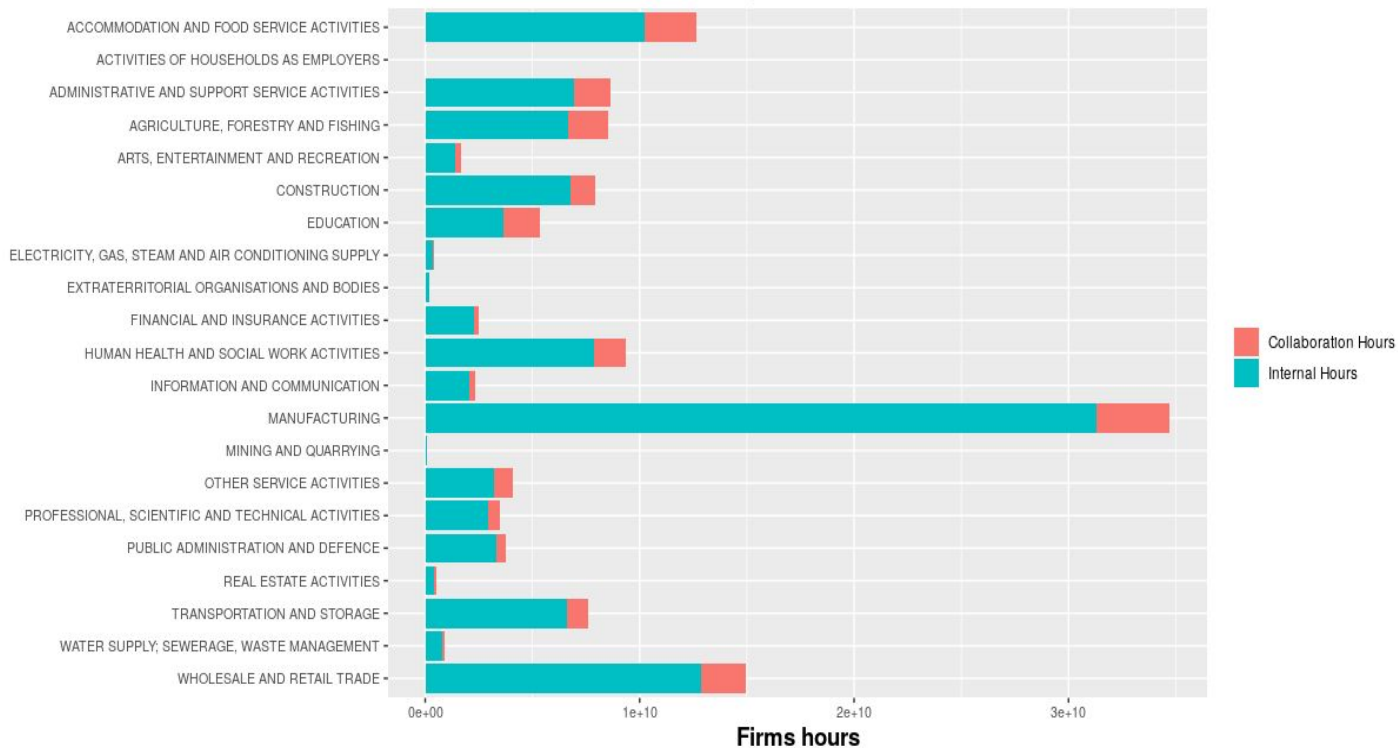

Firms distribution among SLL codes in Friuli-Venezia Giulia

DATASET **HOURS**

There is an unbalance toward manufacturing, that is the 3th for number of firms.

Firms distribution among hours in Friuli-Venezia Giulia

CLEANING **PROCESS**

ATECO
Inexistent or more than one code

HOURS
Lower than a threshold of 1760

ISOLATES
Firms without collaborations

CLEANING
PROCESS

Step by step cleanup process

| OPERATION	ORDER	SIZE
INITIALIZATION	32020	228352
▬ LOOPS	32020	196332
▬ ISOLATES	22616	196332
▬ WRONG ATECO	21131	123020
▬ HOURS < 1760	21017	122720
▬ ISOLATES	19981	122720

02

NETWORK

Exploratory analysis

EXPLORATORY
ANALYSIS

- Full **one-mode** directed network with mutual weighted arcs

- Mutual **dyads** = N/2, others are null

- Density = 0.00031

- **Sparse** network as k ≈ 6n, where k is the number of arcs and n the number of nodes

- Degree distribution is right skewed

- Centrality measures suggest clustering
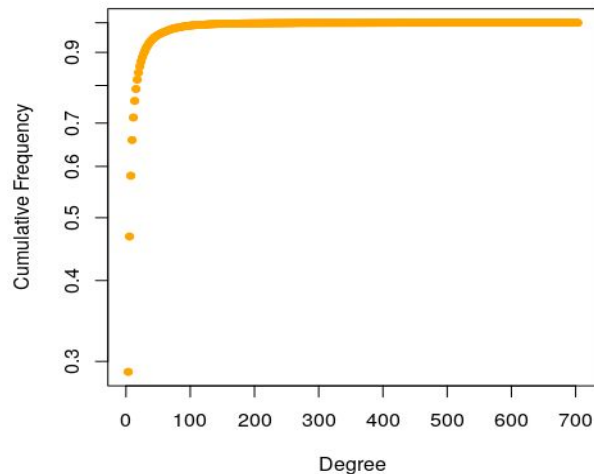
	MEAN	MEDIAN	MIN	MAX
DEGREE	12.28	6	2	704
CLOSENESS	$1.191e^{-05}$	$9.661e^{-06}$	$3.294e^{-08}$	$5.005e^{-05}$
BETWEENNESS	$5.205e^{-05}$	0	0	0.230

SOCIOMATRIX

- **Collaboration** matrix between firms from different province, aggregated by ATECO code.
- Collaboration are **dichotomized**, so ij is the sum of the number of collaborations between companies of ATECO codes i and j.
- Diagonal elements are the number of collaborations between companies of the same category.
- Matrix are **normalized** between 0 and 1 in order to improve readability through the heatmaps.

SOCIOMATRIX **TRIESTE**

Further analysis involved the cell CF, which is the collaboration between manufacturing and construction companies
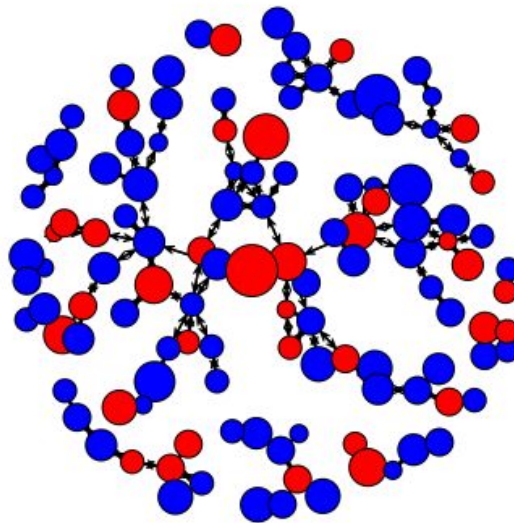

Collaborations in Trieste by ATECO codes
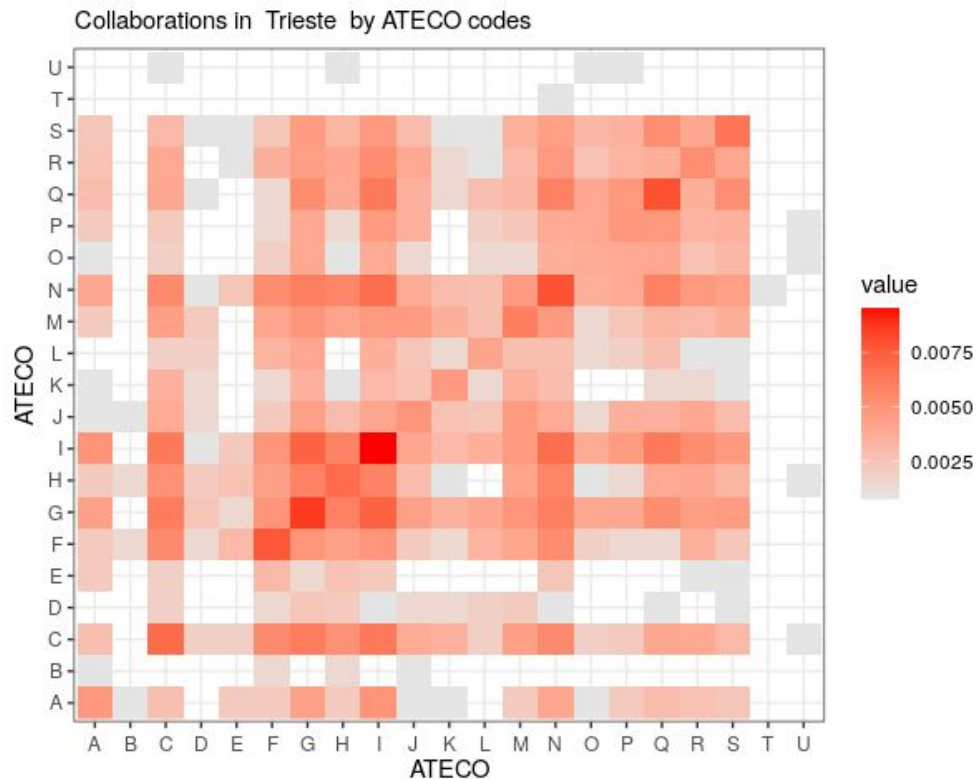
SOCIOMATRIX **UDINE**

Further analysis involved the cell CG, which is the collaboration between manufacturing and retails companies

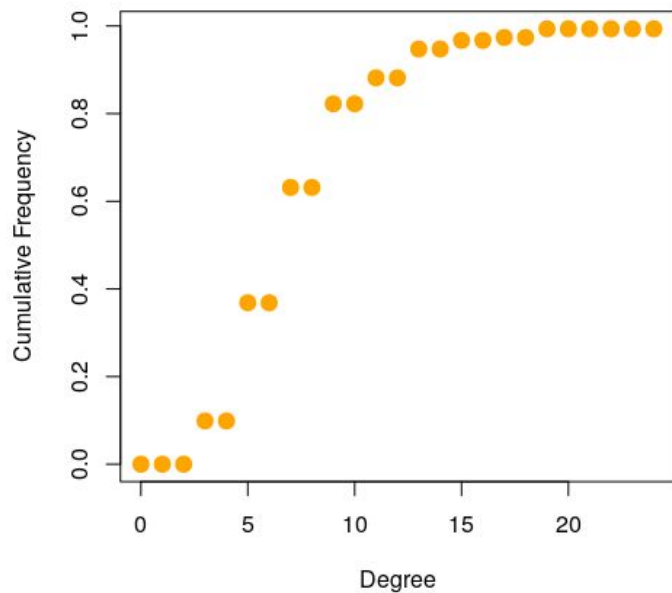Collaborations in Udine by ATECO codes
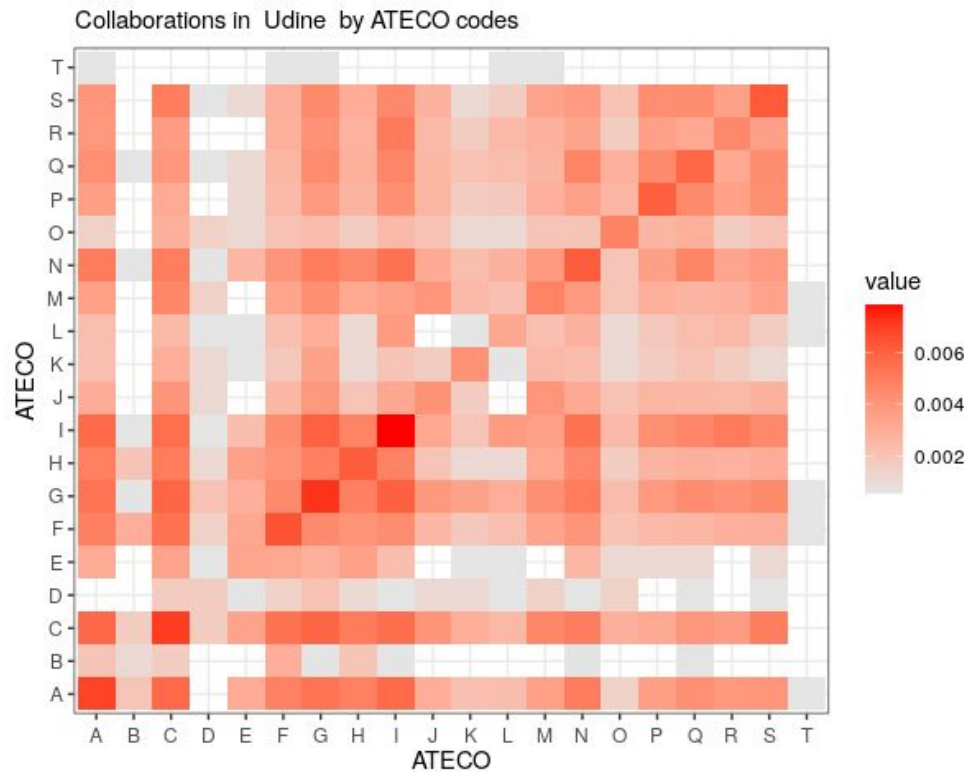
SUBNETWORK **TRIESTE**

- Sparse network

- Size = 506

- Order = 152

- Manufacturing nodes = 47

- Construction nodes = 105

- Edge density = 0.022
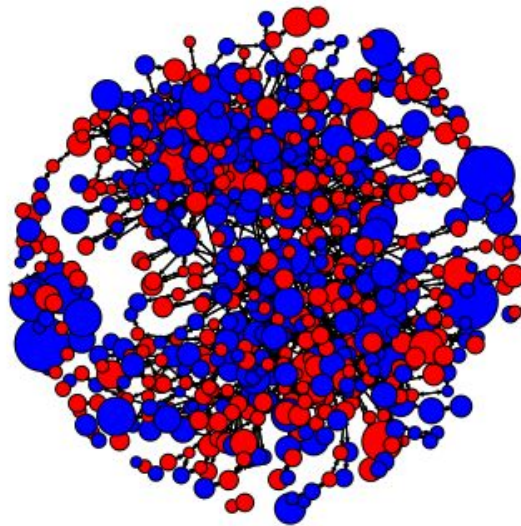

MANUFACTURING 🔴
CONSTRUCTION 🔵

SUBNETWORK
TRIESTE

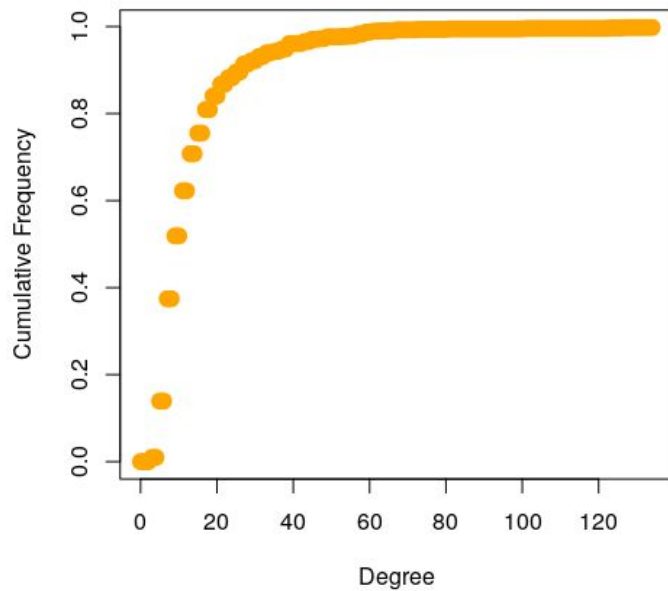|	MEAN	MEDIAN	MIN	MAX
DEGREE	6.658	6	2	24
CLOSENESS	$3.581e^{-06}$	$3.581e^{-06}$	$8.066e^{-07}$	$5.286e^{-06}$
BETWEENNESS	0.049	0.013	0	0.572

SUBNETWORK **UDINE**

- ■ Sparse network

- ■ Size = 5996

- ■ Order = 927

- ■ Manufacturing nodes = 456

- ■ Retail nodes = 471

- ■ Edge density = 0.007

MANUFACTURING ●
RETAIL ●

SUBGRAPH
UDINE

	MEAN	MEDIAN	MIN	MAX
DEGREE	12.94	8	2	134
CLOSENESS	$1.011e^{-05}$	$1.124e^{-05}$	$1.124e^{-05}$	$1.508e^{-05}$
BETWEENNESS	0.011	$3.666e^{-04}$	0	0.366

03

COMMUNITIES

Walktrap and Louvain algorithms

COMMUNITIES **DETECTION**

WALKTRAP

This method runs short **random walks** of few steps and uses the results of these random walks to merge separate communities in a bottom-up manner.
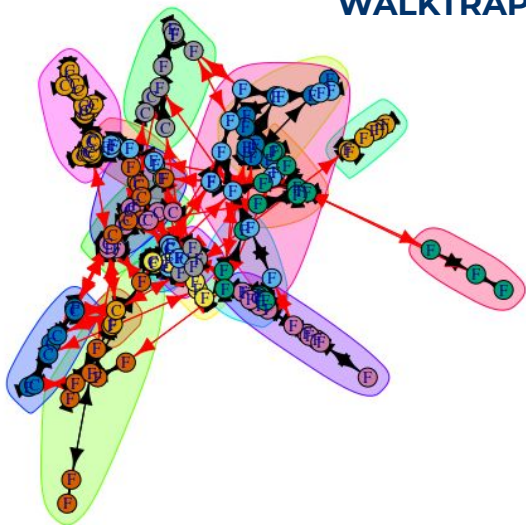
computation time is $O(n^2log(n))$

LOUVAIN

Looks for "small" communities by optimizing **modularity** in a local way, it aggregates nodes of the same community and repeats iteratively until a maximum of modularity is attained.
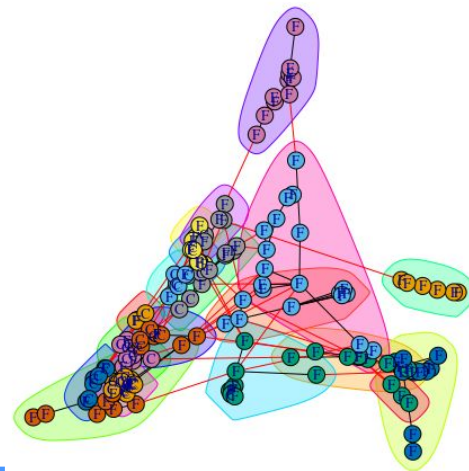
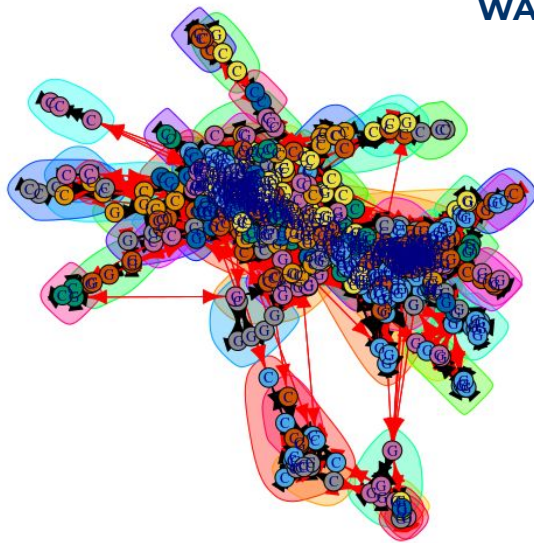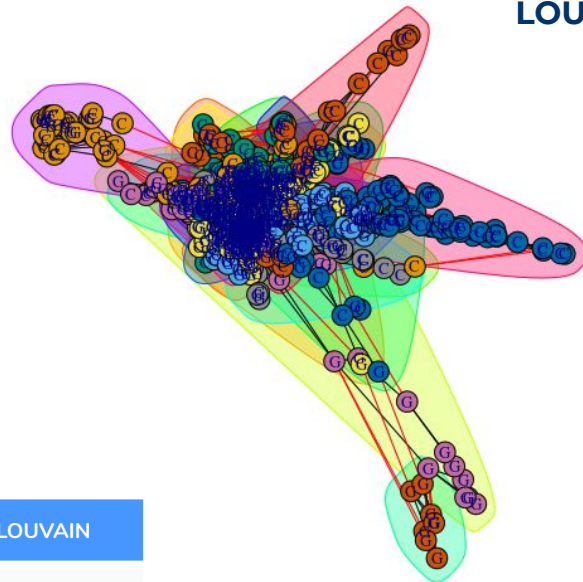computation time is $O(n^2log(n))$

COMMUNITIES **TRIESTE**

WALKTRAP

LOUVAIN

	WALKTRAP	LOUVAIN
MODULARITY	0.848	0.863
# COMMUNITIES	30	19

COMMUNITIES **UDINE**

WALKTRAP

LOUVAIN

	WALKTRAP	LOUVAIN
MODULARITY	0.761	0.799
# COMMUNITIES	142	35

04

ERG MODELS

Simulation on Trieste and Udine subnetworks

ERG MODELS

1. baseline model in which only edges are used as covariate

2. homophily on ATECO codes as factor

3. extracted memberships from the Louvain clustering method as covariate, which should improve at least the intra-cluster edge placement without overfitting

4. Mutual parameter, which reproduce the graphs perfectly as mutuality is 1
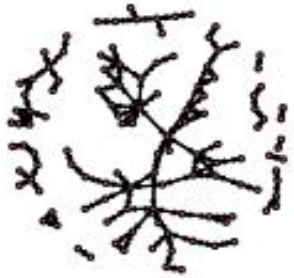
```
nu.01 <- ergm(nu~edges)

nu.02 <- ergm(nu~edges+nodematch("Ateco", diff=T))

nu.03 <- ergm(nu~edges+nodecov("mem"))

nu.04 <- ergm(nu~edges+nodecov("mem")+mutual)
```
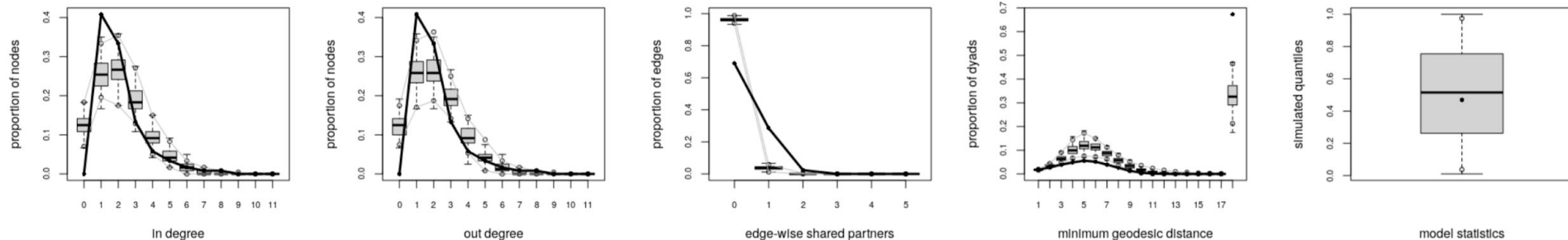
SIMULATIONS TRIESTE

	SIMULATION			
#	**1**	**2**	**3**	**4**
AIC	2536	2539	2537	1835

SIMULATIONS UDINE

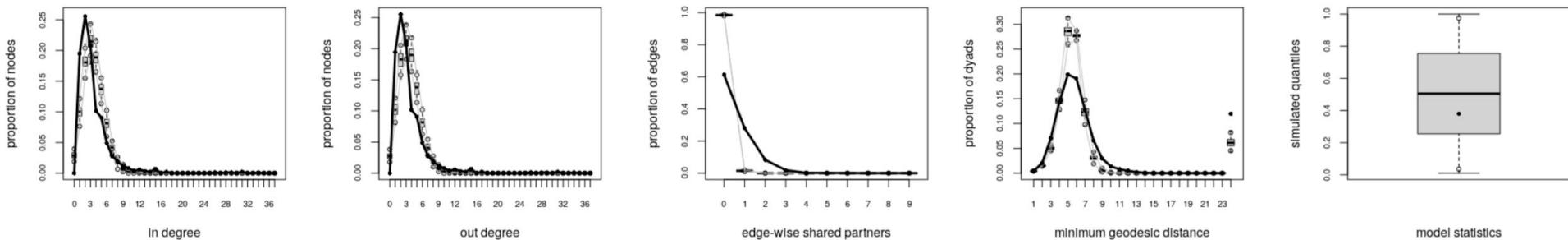|	SIMULATION			
#	1	2	3	4
AIC	34281	34283	34285	27569

GOF TRIESTE
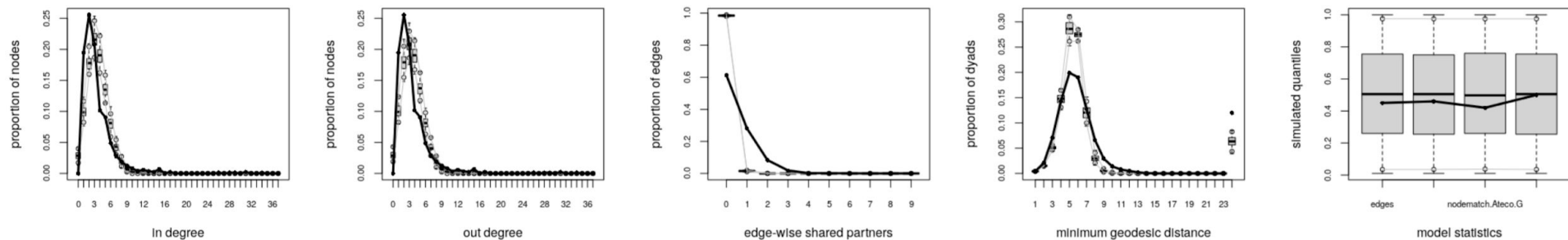
EDGES

EDGES, LOUVAIN MEMBERSHIP
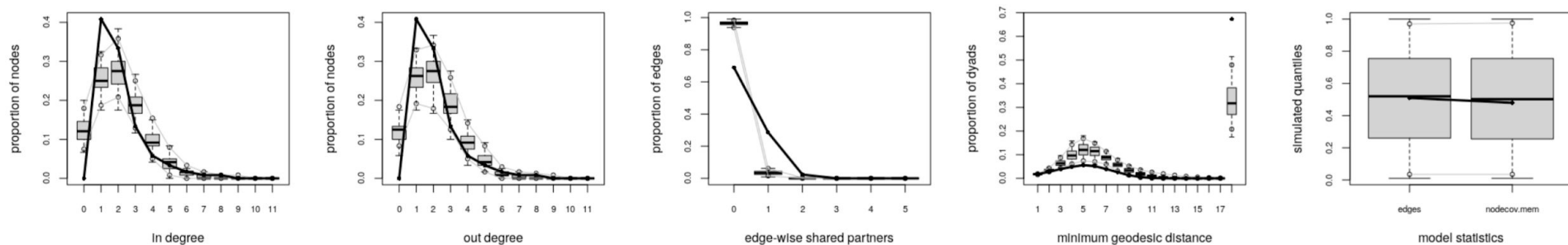
GOF UDINE

EDGES

EDGES, ATECO HOMOPHILY, LOUVAIN MEMBERSHIP

CONCLUSION

This work shows an approach to understand the relationships between firms of different sectors through their collaboration

The dataset is huge with a lot of information, but also incomplete and should be treated carefully.

Firms tends to collaborate in the same sector as shown in the sociomatrices

Community detection showed good results and a more thorough analysis can expose underlying patterns in companies partnerships

THANK YOU