UNIVERSITÀ DEGLI STUDI DI MILANO
UNIVERSITÀ DEGLI STUDI DI BRESCIA
UNIVERSITÀ DEGLI STUDI DEL PIEMONTE ORIENTALE

PhD PROGRAM
ECONOMIC SOCIOLOGY AND LABOUR STUDIES – 29th cohort

DOCTORAL THESIS

Internal segmentation, residential patterns and job-related spatial mobility for independent professionals: An original analysis on primary data.

SPS/09, SPS/07, IUS/07, SECS-P/07, SECS-P/10, SECS-S/04, M-PSI/06

PhD Candidate
CLAUDIA COLETTA

Supervisor
Prof. PAOLO PERULLI

PhD Director
Prof. PROF. GABRIELE BALLARINO

ACADEMIC YEAR
2017/2018

The PhD program Economic Sociology and Labour Studies (ESLS) (29th cohort) stems from the collaboration of four Universities, namely Università degli Studi di Brescia, Università degli Studi di Milano and Università degli Studi del Piemonte Orientale "Amedeo Avogadro". The University of Milan serves as the administrative headquarters and provides the facilities for most teaching activities.

“ Believe those who are seeking the truth. Doubt those who find it ”

André Gide (1869-1951)

Contents

List of figures

List of tables

Introduction

The aim of this research is mainly to let readers focus their attention on the emerging of a brand new category of workers in the labour market: the Independent professionals (IPros)[1].

IPros are highly skilled, independent self-employed individuals who work for themselves and do not employ others; they are what we usually know under the name of "freelancers".

Over the past 30 years the change of working conditions, the tertiarization of the labour market and the rising of the so-called knowledge economy have brought to the fore new forms of "a-typical work".

There have been several studies about atypical forms of employment, whereas new professionals (IPros) have been neglected and mostly disregarded by socio-economic literature (Ranci 2012) despite their rising numbers.

Data from European Labour Force Survey (ELFS) 2011 show a steep increase of own-account workers in the last 10 years. Is this trend just a reflection of temporary recessionary conditions in which people can't get jobs?

We believe this to be more than a temporary phenomenon, as witnessed by many proto trade unions which have been born in the last decades (the European forum of Independent Professionals in 2010, the American Freelancers Union in 2003).

The grounding idea of this research is to contribute in exploring this new phenomenon by mapping living places, tracking mobility patterns in relation to socio-demographic characteristics of this heterogeneous group of workers which Bologna and Fumagalli (1997) defined as "Second generation independent workers".

Beside the accepted idea that labor flexibility is mostly related to time flexibility, we want to stress the role of the spatial dimension: labor flexibility often involves spatial flexibility, too. This is why the approach adopted mainly focuses on living places and mobility patterns of IPros, where causal explanations are found both at micro and macro levels: socio-demographic variables as well as infrastructural ones (land use)[2].

[1] The label Independent professionals was first used by The Association of Independent Professionals and the Self-Employed (IPSE), a British not for profit organisation established in 1999 and which now covers interests of freelance consultants and contractors as a representative body for freelancing in the UK.

[2] When we talk about mobility we cannot avoid talking about local infrastructure network supporting and making possible the flow of people and the flow of informations and data.

This dissertation explores the topic by adopting an Interdisciplinary perspective: it brings together contributes from different literature fields such as economic geography, urban sociology, labour studies and mobility studies. Literature references from different disciplines are accurately reviewed in Chapter 1.

The method of study is based on quantitative analysis mainly performed on survey data, due to the lack of valid data sources in Italy.

Specifically Chapter 2 presents a detailed description of survey design and data collection process, whereas Chapter 3 specifies methodological strategies applied for data cleaning, coding and editing and introduces descriptive results related to research questions.

GIS (Geographical Information System) tools have been related to the survey in order to provide a full extensive spatial description of social phenomena investigated: Chapter 4 is based on GIS and spatial statistics techniques by using data collected at micro and macro levels.

To sum it up here it is our challenge: studying a new subject, the Independent professionals, from an unusual[3] perspective in sociology: the geographical space.

[3] Sociological studies on panel data are much more diffused than studies on spatial ones. We, as sociologists, often tend to forget that phenomena are not just time based but they are spacially based, too.

Chapter 1

Literature review and Research design

1.1. Economic sociology and labor studies

1.1.1 *Macro-economic changes and the labor market: the theoretical background*

From the early 1970s the features of economic production have changed in many ways thanks to

- the development of new technologies for information and communication[1]
- the rising tertiarization of the economy
- the emerging of new economic sectors (finance, ICTs, management consultancy, brokerage services)
- the network model in the production chain structure

The hierarchical Fordist system has been replaced by a sobstitution-oriented market[2] which requires a higher degree of flexibility[2]. This flexibility has led to just-in-time and small scale production, relocation of manual routinized industrial activities and, on the other hand, outsourcing and subcontracting for highly specialized ones (Reyneri 2011).

The organisation of work has changed considerably over the last two decades, with new forms of work having gained in importance: the whole economy has moved from the dominant hierarchical model of long-term labour to a bunch of different contractual positions which are usually brought back to unity under the comprehensive label of "a-typical work"[3].

[1] ICTs has deeply modified space-time coordinates for the production cycle and has led towards the knowledge society by allowing the storing and the analysis of so-called "big data" and their free circulation in the form of "open data". Sabel and Piore (1984) have been among the first to remark the shift from the Fordist standardized mass-production model to systems of flexible specialization. Later Andersen and Regini (2000) have been able to identify 3 different models of competitive production: diversified quality production, flexible specialization and flexible mass production (or Toyotism).

[2] Flexibility mainly concerns the number of workers (numerical flexibility), their tasks (functional flexibility), their wage differentials (wage flexibility) and their working times (time flexibility).

[3] Atypical (or non-standard) work refers to work relationships beyond full-time and permanent employment with full social insurance coverage, showing clear features of precariousness. The term usually includes part-time work, marginal work, temporary work, agency work and dependent forms of self-employment.

Moreover the production has shift towards the so-called "Knowledge economy"[4], the new postFordist economic setting[5] aiming at providing services and intangible goods as the main source for global competitiveness.

The post-industrial transition has been followed by a productive fragmentation facing a stronger international competition on the global trade market. The "pulverization" (Ranci 2012: 22) of the tertiary sector has witnessed an increase in the number of atypical workers and independent professional workers in new economic sectors.

As a result, the traditional threshold between dependent and independent workers has become blurry and internal segmentations has structured the category of professionals[6].

On one side we find the brand new «economically dependent self-employed»; on the other we find the new professional self-employed having no formal recognition of its expertise and educational career, facing risks related to job and income instability, time pressure, pressure to perform, or a problematic work-life balance.

Within this scenario the flexible worker is the new emerging social actor of the post-modernity.

New difficulties emerge when defining which social class these workers belong to. According to Goldthorpe's scheme (1992) the petite bourgeoisie of own-account workers was a class itself marked by a relatively high income and lower educational profile; now the "status incongruity" is overturned: the new professionals in the knowledge economy face income instability despite having high educational profiles (Ranci 2012: 59).

1.1.2 Micro perspective: the Indipendent Professional as the object of the study

Self-employed workers in advanced services (finance, counselling, producing-distributing information, auditing and so on) have increased their share in employment in most countries, and they display the highest growth in employment rates in the leading metropolitan areas of the world.

[4] The goals of the knowledge economy have been addressed by European Commission in the strategy Europe 2020 (which followed Lisbon Strategy 2000-2010).

[5] Here they take place new forms of social inequalities mainly based on urban form of agglomeration and segmentation among those accessing "Apple jobs" and those providing lower class services in the "Big-mac jobs" (Giddens 2007).

[6] A.J. Scott (2011: 5) talks about "new divisions of labor that are appearing in the detailed organization of production and in related processes of social re-stratification".

Data from European Labour Force Survey (ELFS) 2011 show that among self-employed the share of those working without employees exceeds the share of self-employed persons with employees in all Member States.

Table 1. Own-account workers in the European Union in per cent of employment (15-64 years), 2011(%)

Country	*Self-employed with employees %*	*Self-employed with employees (1000)*	*Own-account workers %*	*Own-account workers (1000)*	*Total %*
EU-27	4.2	9.006.1	10.2	21.572.3	15.8
Austria	4.8	193.7	6.6	267.0	12.7
Belgium	4.1	183.0	8.7	389.9	13.8
Bulgaria	3.6	105.6	7.3	211.7	11.9
Cyprus	4.5	16.6	10.6	40.4	16.4
Czech Republic	3.5	169.9	13.7	655.7	17.7
Denmark	3.6	93.9	4.8	127.1	8.6
Estonia	3.9	23.6	4.2	24.7	8.0
Finland	4.0	96.7	8.2	200.2	12.5
France	4.4	1.130.5	6.5	1.653.4	11.3
Germany	4.5	1.772.8	6.0	2.328.5	10.9
Greece	7.6	297.8	22.8	897.6	35.6
Hungary	5.2	196.2	6.2	232.9	11.8
Ireland	4.7	83.3	10.3	182.7	15.6
Italy	6.3	1.427.0	16.2	3.648.1	24.0
Latvia	3.7	30.9	6.5	54.0	11.2
Lithuania	2.4	29.0	6.6	81.3	10.3
Luxembourg	2.6	5.7	5.1	11.4	8.3
Malta	4.1	6.8	8.9	14.6	13.0
Netherlands	3.8	311.8	9.9	817.0	14.3
Poland	4.1	637.4	14.4	2.226.2	22.0
Portugal	5.0	222.3	11.6	524.6	17.1
Romania	1.2	105.6	16.7	1.459.0	29.8
Slovakia	3.5	80.5	12.3	284.3	15.9
Slovenia	3.6	32.8	8.3	75.9	15.2
Spain	5.0	913.4	10.4	1.892.9	16.2

Sweden	3.6	159.8	5.8	259.3	9.5
United Kingdom	2.4	679.6	10.7	3.012.0	13.3

Source: Eurostat, LFS 2011 (http://appsso.eurostat.ec.europa.eu/nui/submitViewTableAction.do)

Beyond tax policies and institutional incentives, the numerical relevance of the phenomenon of the revival of self-employment[7] (both in absolute and relative terms) could be explained by several macro-factors, some of them we already mentioned:

- The sectoral shift of the labour force from manufacturing to knowledge-intensive services offer new possibilities for self-employed work. Especially to become own-account worker entry barriers are lower and the lack of capital is no more seen as an obstacle. "Specialized expertise, reputation and educational and professional qualifications" are the new prerequisites (Bryson et al. 1997: 345)

- The educational inflation and the progressive specialization of knowledge

- The outsourcing strategy and the stress towards numerical flexibility (Luber and Leicht 2010) is the basis of the "unemployment push theory" (Bögenhold and Staber 1990). Self-employment becomes a strategy for reintegration into the labour market[8].

- Self-employment among immigrant ethnic minorities is the road to higher economic returns with respect to wage earning (Clarke and Drinkwater 2000)

Among self-employed own-account workers we find a specific category, the Independent Professionals (Ipros). The term is used together with those of "Second generation independent workers" [9] and "Knowledge workers".

They are synonymous generally used to refer to a very wide group of workers including researchers, professionals belonging to professional associations, experts in new emergent economic fields not yet formally recognized as professions.

[7] The growth first started during 1980s with new professions entering the market (Luber and Leicht 2010).

[8] This theory is a valid explanation to high number of own account workers recorded in Greece, Italy, Romania. In these countries such high figures are also due to the phenomenon of the so-called bogus self-employment, which is another face of precariat.

[9] This definition, provided by Bologna and Fumagalli (1997), remarks the deep diversity with respect to traditional forms of independent work such as the craftsman, the trader or the professional belonging to a professional association (i.e. medician, lawyer, business consultant, notary etc.).

For the purpose of this research we will consider by the name Ipros a very specific target of workers with the following characteristics:

1. working for themselves with no employees
2. working in advanced service activities
3. being highly skilled and well qualified (tertiary education)
4. not being enrolled in any formal professional association

The European Labour Force Survey provide us a first rough estimate of the numerical consistency of these workers.

Table 2. Own-account workers 20-64 years by occupation ISCO08 Professionals + technicians and associate professionals, 2011, (% with respect to active population aged 20-64)

GEO/TIME	*2005*	*2006*	*2007*	*2008*	*2009*	*2010*	*2011*	*2012*	*2013*
UE-27	2.51	2.62	2.70	2.70	2.70	2.82	2.97	3.05	3.05
Belgium	3.13	3.11	3.09	3.08	3.15	3.25	3.30	3.51	3.67
Bulgaria	0.95	0.89	0.90	1.03	1.06	1.07	0.95	1.07	1.18
Czech Rep.	3.55	3.67	4.11	4.22	4.13	4.28	4.28	4.43	4.16
Denmark	1.16	1.25	1.37	1.36	1.58	1.71	2.02	2.24	2.22
Germany	2.74	2.79	2.77	2.78	2.81	2.94	3.19	3.27	3.09
Estonia	0.69	0.96	0.98	0.85	0.81	1.13	1.14	1.20	1.31
Ireland	1.79	1.84	1.87	1.90	2.02	2.12	2.31	2.21	2.25
Greece	3.04	3.23	3.24	3.31	3.30	3.36	3.68	3.65	3.53
Spain	2.19	2.29	2.38	2.31	1.93	1.93	2.05	2.11	2.25
France	1.25	1.43	1.48	1.40	1.48	1.66	2.14	2.07	2.10
Croatia	0.73	0.67	0.90	0.85	0.92	0.83	0.76	0.96	1.12
Italy	5.91	6.11	6.33	6.16	6.00	6.19	6.18	6.15	6.01
Cyprus	2.63	2.43	2.29	2.55	2.50	2.62	2.62	2.77	2.82
Latvia	.	0.91	0.92	0.99	1.27	0.99	1.14	1.21	1.29
Lithuania	.	.	.	.	.	.	.	0.79	0.95
Luxembourg	2.00	2.03	2.11	1.95	2.38	2.26	2.59	3.16	3.07
Hungary	1.56	1.42	1.58	1.60	1.54	1.42	1.72	1.59	1.57
Malta	1.26	1.50	1.53	1.69	1.73	1.88	1.43	1.57	1.74
Netherlands	3.28	3.31	3.53	3.89	3.91	4.35	4.58	4.85	5.12
Austria	2.22	2.30	2.61	2.68	2.81	2.81	2.60	2.64	2.94
Poland	1.36	1.43	1.68	1.74	1.92	2.04	2.02	2.17	2.30

Portugal	1.28	1.38	1.40	1.35	1.55	1.47	1.56	1.70	1.84
Romania	0.22	0.29	0.30	0.28	0.32	0.41	0.45	0.50	0.50
Slovenia	0.74	0.89	0.81	0.85	1.14	1.30	1.90	1.66	2.01
Slovakia	2.29	2.28	2.38	2.66	2.98	2.93	2.93	3.00	2.90
Finland	1.74	1.80	1.70	1.94	1.98	2.04	2.42	2.53	2.53
Sweden	2.14	2.26	2.32	2.32	2.39	2.44	2.37	2.37	2.39
U. K.	2.83	2.99	2.99	2.99	3.08	3.17	3.26	3.42	3.42

Source: Eurostat, LFS 2011 (lfsa_esgais, own elaboration)

(http://ec.europa.eu/eurostat/data/database)

The trend toward an incremental growth of Ipros in Europe is visible and it raises the urgency to study this new phenomenon taking place in the labor market.

Finally, what is the socio-economic profile of the individuals we are going to study?

Generally speaking "knowledge workers" are not a homogeneous group. This is quite intuitable especially when looking at the different labels used to refer to them; "knowledge workers (Drucker 1993), "freelancers", "second generation autonomous workers" (Bologna 1997), "creative class" (Florida 2003). Mainly we could distinguish 3 sub-categories among them: 1) executives, managers and entrepreneurs 2) highly specialized professionals 3) higher level technicians[10].

We will consider only the group 2) and 3) which mostly reflect the requisite of being own-account workers[11].

They basically are the one showing on average a high level of status incongruity.

They take part to immaterial processes (inventing, researching, managing, controlling, designing, communicating etc.) and they have to have hard and soft skills (leadership, team-work abilities, risk-taking, problem-solving attitude).

[10] Among them are included i) industrial technicians, engineering professions, environmental security professions, computer scientists ii) professions in the welfare system (teachers, nurses, social assistants, physiotherapist, etc.) iii) accountants, insurance agents, bankers, real estate brokers, financial brokers and so on. In the Italian case they all together form the bigger group working for public and private organizations (Consorzio AAster 2011, research report).

[11] "The post-Fordist technician is less and less a worker belonging to an organization" (Pichierri 1997); identity is mediated by the development of professional qualities. This gradual transformation is mainly due to the changing of job contents and the organizational setting (lean manufacturing, networked enterprises); he/she works for big-medium firms ("transational firms") which often use outsourcing (Perulli 2012: 8).

But despite being full of knowledge resources, these workers are frail. They undergo several risks concerning long-term inactivity and underemployment, losing creativity and inventiveness, lack of professional identity and isolation[12].

They could be defined as the precarious[13] of the new knowledge economy, since working is no more a stable activity but it splits into a continuous flow of events marked by changes and uncertainty[14].

Flexibility in the working environment has made it difficult to the individual to make long-term commitments; disorientation and fragmentation of coherent objectives for the individual is the consequence of job insecurity.

Differing from traditional professionals, Ipros have not access to the path towards collective action, and cannot ask for any form of protection from social dumping.

Thus, despite they work in new economic leading sectors, they rarely are part of the managerial elite that occupies the leading positions in our societies. When we talk about Ipros we do not talk about a uniform rampant emerging class of businessman.

The Idea of internal segmentation within Ipros includes some socio-demographic variables such as age and gender.

The following table shows clearly the age and gender divide for the Italian case.

[12] Using Sennett's words (1998) we would talk about "the corrosion of character". Flexibility, flextime, teamwork, delayering and ever-changing working conditions create new forms of oppression ultimately disorienting individuals and undermining their emotional and psychological well-being. It is a restlessness space where success and fragility intertwine together.

[13] Rodgers (1989) explains precarious work along four dimensions: (i) the degree of certainty of continuing employment; (ii) the degree of control over working conditions, wages and the pace of work; (iii) the degree of labour protection (e.g. against discrimination, unfair dismissal, unacceptable working conditions, social protection); and (iv) the income level.

Table 3. Average income for Ipros in Italy (€), 2012-2011

Age class	*Sex*	*Average income 2011*	*Average income 2012*	*Difference 2012-2011 (abs. value)*	*Difference 2012-2011 (%)*
20-24	M	9698	9032	-666	-6,9
	F	9835	9432	-403	-4,1
	TOT	9759	9210	-549	-5,6
25-39	M	12762	11742	-1020	-8,0
	F	11799	11287	-512	-4,3
	TOT	12288	11515	-773	-6,3
30-34	M	15535	14890	-645	-4,1
	F	12694	12158	-536	-4,2
	TOT	14248	12158	-536	-4,2
35-39	M	19393	18619	-774	-4,0
	F	14362	13831	-531	-3,7
	TOT	17202	16500	-702	-4,1
40-44	M	22577	22203	-373	-1,6
	F	16085	15664	-420	-2,6
	TOT	19778	19368	-410	-2,1
45-49	M	24237	23959	-279	-1,1
	F	17499	17071	-428	-2,4
	TOT	21490	21060	-431	-2,0
50-54	M	24770	24616	-154	-0,6
	F	17525	17354	-172	-1,0
	TOT	22170	21930	-240	-1,1
55-59	M	24642	24357	-285	-1,1
	F	18076	17871	-205	-1,1
	TOT	22789	22432	-358	-1,5

60-64	M	24631	24484	-147	-0,6
	F	18538	18038	-500	-2,7
	TOT	23336	23056	-280	-1,2
65-69	M	24522	24804	282	1,1
	F	17656	17392	-264	-1,5
	TOT	23349	23482	134	0,6
70 and over	M	22058	22071	13	0,0

Source: own elaboration on administrative data from Gestione Separata INPS (www.inps.it/webidentity/banchedatistatistiche)

The data show 2 cleavages related to age and gender.

Professionals up to 39 years old earn much less than older ones; in addition when comparing data between 2011 and 2012 the negative difference in average income is no lower than 4% for younger workers, whereas it seldom exceeds 2% for older ones.

On the other hand the gender issue related to income becomes relevant for workers over 30; the negative difference raises from around 18% (30-34 age class) up to almost 30% (65-69 age class).

1.2. Urban studies and Economic geography

1.2.1 From urban sociology to the space of flows

The transformation from Fordist to Post-Fordist, from the industrial to the post-industrial city or from modernity to post-modernity affected all cities deeply.
Since the beginning of the 1980s, together with the permeation of neoliberal ideologies and practices, the urban phenomenon has come to the fore; the city has emerged once more[14] as object of research itself as well as the reference point for studying social economic and political processes. Social polarisation, growing inequality, increasing residential segregation have been core issues of urban sociology for many years.

Still today, far from being dissolved by globalization processes, cities are places where modern economic productive systems settle their roots. The coming of two technological macro systems, such as the one affecting physical mobility and the other improving information transmission, has contributed to the radical change of urban morphology.

ICTs revolution was indeed supposed to provoke the decline of dense urban forms, and to diminish spatially localized social interactions. Nevertheless that was not the case: urban areas are imploding and exploding at the same time.
This dialectical process, recalled by Brenner's latest book (2014), is an essential analytical tool for any theory concerning urbanization in the twenty-first century. Concentration and agglomeration take place in parallel with "the extension of urban fabric, intensification of interspatial connectivity across places" (p.197). Using Le Febvre's words (1970: 14) we should look at both "the tremendous concentration (of people, activities, wealth, goods, objects, instruments, means, and thought) of urban reality and the immense explosion, the projection of numerous, disjunct fragments (periphery, suburbs, vacation homes, satellite towns)"

The explosion of urbanity is typical of contemporary urban societies which are highly, and intrinsically mobile. Urban lifestyles and business processes are dispersed spatially and thus increasingly dependent on extensive flows of persons, goods and information.

Castells (1996) analyses the transformation of localization patterns by developing the idea of "space of flows". More precisely this term addresses the process of connecting services and economic actors with different intensity and at a different scale, depending upon the relative importance of activities located in each area. "The global city is not a place, but a process. A process by which centres of production and

[14] First came studies by Max Weber (1920-21) and Simmel (1903). Later during the 1920s and 1930s the development of the so-called "Chicago School", whose main theorist was Robert Park, specialized in research in the urban environment through ethnographic fieldwork (human ecology approach).

consumption of advanced services are connected in a global network, while simultaneously downplaying the linkages with their hinterlands, on the basis of information flows" (Castells: 417).

The "infinite city"[15] (Bonomi and Abruzzese 2004) or "megalopolis" (Gottmann 1961) is on the foreground; the spreading of economic activities and functions on the territory leads to a new urban gigantism[16] and, on the other hand, to a parallel de-urbanization.

For a long time urbanity has been defined in terms of dots, lines and boundaries. Today this description is meaningless; urban realities are local and global at the same time and they are continuously crossed by flows of information, people and money. Cities are no more flat points on the map. They are assemblies of economic relations (at long or short distance) with different intensities in different places (Amin, Thrift 2001: 82).

To conclude, there is the need for a new spatiality beyond old territorial stereotypes. We should talk much more about "urban environment" and much less about "cities". The "cityness" (Brenner 2010) is an evolving condition set at different scales; it is not a specific object territorially defined. It is for this reason that spatial scales and territories are "one of the most daunting methodological challenges facing contemporary social science" (Brenner 1998: 28).

1.2.2 The polycentric city region within the network-based society

After addressing the issue of explosion and the idea of a diffused "cityness", we should look back to the other side of the coin: tendences towards rising agglomeration[17].

The complementarity of global and local is pointed out by the fact that long-distance networks have not diminished the importance of local ones; indeed global nodes[18] are part of particular specific places which are the basis for local competitive advantage in location choices. Physical proximity still matters, once again.

Capitalist economies of agglomeration has been built on the advantages (lower transaction costs, availability of specific services and know-how) deriving from spatial proximity, despite during the 1970s

[15] According to Nancy (1999) "the centre is everywhere and the circumference is nowhere, or the other way round: the city is a dispersed totality".

[16] In the work of Soja (2000) the term "exopolis" refers to the growth of the outer city, suggesting the increasing importance of exogenous forces.

[17] Scott e Storper (2003) underline the urgency to go against theoretical approaches that suggest de-centralization and spread to be the only output of globalization.

[18] These central urban nodes operating at world-scale economy are commonly defined as "World Cities" or "Global Cities" (Sassen 1994).

the de-industrialization process led to the relocation of many economic activities far away from the local market[19] and the rule of proximity was discredited. The stress was on the role played by new Information technologies (and dematerialization of most economic activities) in creating a global space for interacting urban nodes.

Only later during the 1990s there has been a rediscovery of the territory as a relevant variable for economic processes: economies of scale still play a relevant role and local agglomeration trends raise the economic competitiveness of specific territories (Porter 1995); in other words urban local market still has a significant weight.

This way we have drawn the path from relocation to new forms of territorialisation such as the so-called "global city region", the nest for global competitiveness which has its roots in the post-industrial knowledge economy.

The term "Global City-Regions" was first introduced in the academic debate by Allen J. Scott in the 1990s. Scott (2001) provides the following definition: "a global city-region can be said to comprise any major metropolitan area or any contiguous set of metropolitan areas together with a surrounding hinterland of variable extent whose internal economic and political affairs are bound up in intricate ways in intensifying and far-flung extra-national relationships".

Whereas the term "global cities" (Sassen 2001) mainly focuses on some central nodes being part of a global network of external relations, the term region stresses the relations developed among nodes within the local system (not outside it). It is a different scale perspective.

Against the rhetoric of a-spatiality pointing at territorial units being dissolved away as definite geographic entities by processes of globalization, city-regions are actually becoming increasingly central to the conduct and coordination of modern life. Global inputs do produce local outcomes: dense urban agglomerations increase in size, mainly due to recent improvement in technology of transportation[20].

Accordingly, the idea of "regionalization of the urban" has gained ground and the term "World Cities" is used no more with reference to big monocentric metropolis, rather to dense and spread polycentric agglomerates, such as Kansai region in Japan, the Rhine-Ruhr in Germany or the Randstad region in the Netherlands (Friedmann, 1995). The resurgence of regions has thus manifested through the formation of

[19] Production activities undergo decentralization, whereas the executive core remains strongly centralized.

[20] Key elements of the city region are transports and local infrastructures, allowing for the increasing number of immaterial flows (i.e. optical fibre and internet cables) as well as long (i.e. international airports, ports) and short distance material ones (i.e. high-speed trains, highways etc.).

new territorial units, going beyond traditional administrative borders: the global city regions, dense locational clusters especially for new types of leading economic activities[21].

As Sabel noticed (1998), it is the long-standing weakening of the economic role of the national state which gave an important contribute to the rising of this new regional perspective.

Brenner (1999) lists globalization and transition to post-fordism as the background phenomena which induced processes of rescaling on the territory together with the decline of national states; some of the regulatory functions carried out by central states have been drifting towards higher or lower levels of spatial resolution[22].

Moreover the "Neo-regionalism" has favoured the strengthening of political powers of local actors within administrative boundaries[23]. The man issue is that there is a mismatch between these entities: economic actors such as "city regions" [24] and political ones such as administrative regions.

Globalization raises many new questions about economic regulation or governance at all spatial levels: the model of the city-region could be an attempt to answer at least part of them. To summarize: "many regions are now faced with the choice of either passive subjection to external cross-border pressures, or active institution-building, policy-making, and outreach in an effort to turn globalization as far as possible to their advantage. Regions that take the latter course are likely to find themselves also faced with many new tasks of political coordination and representation" (Scott 2001).

The governance of urban regions is an open issue; they are informal territorial scales and lack a formal regulation, mostly being the field of play for private for profit actors. Garavaglia (in Perulli 2012) highlights the role of the state or pioneering initiatives, built up by regions and associations representing interests of workers, in promoting the diffusion of technology, knowledge, training, the formation of consortia for internationalization of the area.

[21] Economic activities like high-technology industry, neo-artisanal manu-facturing, cultural-products sectors, the media, business and financial services, and so on. "Advanced services, including finance, insurance, real estate, consulting, legal services, advertising, design, marketing, public relations, security, information gathering, and management of information systems, but also R&D and scientific innovation, are at the core of all economic processes" (Castells 1996: 416).

[22] Scott (2001) recognises a new spatial hierarchy of political-economic-social relations, featuring 3 levels: a global level (mainly represented by international bodies such as the European Union), a declining national level and a rising regional one.

[23] In Italy the constitutional reform of title V conceded new political autonomy to Regions and the law 328/2000 established the decentralization of services.

[24] Several authors (such as Herrschel and Newman, 2002) identify city regions such as those areas being the key arenas for economic growth.

Finally, to complete the picture, we must add the concept of polycentrism, a vague and polysemantic one. It refers to the existence of several urban centers, similar in size and close to each other[25] (Meijers, 2007). The functional dimension is fundamental: these nodes are specialized in complementary economic sectors and mostly cooperate in order to make the whole regional area competitive in the global arena. First studies on the topic (Dieleman and Faludi, 1998; Kloosterman and Mustard, 2001) mainly focused on polycentric region in the North Western Europe: Netherlands, Belgium, Germany.

Notwithstanding their goal of improving economies of scale while avoiding diseconomies of agglomeration (Bailey e Turok 2001), the role of big and small urban nodes in the same geographical area forming a polycentric urban region remains controversial. As Cappellin (1991, cited by Castells 1996) underlines: "The relative importance of the city-region relationships seems to decrease with respect to the importance of the relationships which interlink various cities of different regions and countries (…) New activities concentrate in particular poles and that implies an increase of disparities between the urban poles and their respective hinterlands".

This statement points out disparities which affect smaller cities, subject to the predominance of metropolis. However if we look at the Italian case as an original example of city region we could rightly define it as an example of fair polycentrism, as many studies point out (Perulli 2012, Perulli and Pichierri 2010).

The role played by Milan as a global city does not prevent other second-level nodes from weaving a dense relational network; a part from some very specific and rare functions, each city is able to locally provide knowledge intensive activities in the service sector (Garavaglia in Perulli 2012).

No more a hierarchical level (Chinese boxes model), but an horizontal one, where the scale of the city (medium, small, large) does not tell us much about the system of flows and mobility, which are the main explanatory factor for local development (Perulli 2012: 27).

The internal variety of city-region is relevant in making them competitive "transactional spaces" where production chains, innovation processes and labour markets get into gear. The process is facilited by the increasing number of advanced infrastructure connecting the nodes and allowing higher mobility to knowledge workers. Indeed people themselves, besides businesses, may produce consistent flows with very dense networks.

Medium-size cities are relevant outposts for the emerging of these networks; they are the place where flows enter and exit and it is not by chance that the number of Italian middle cities has grown from 491

[25] A different interpretation recognises the presence of polycentric metropolis being the biggest nodes within a cityregion made up by 10 up to 50 urban actors, physically separated but part of a functional network built up around bigger nodes (Hall and Pain 2006).

to 1054 between 1911 and 2011 (Pennati and Garavaglia 2014) gathering together along territorial corridors which connect main centres to each-other[26].

1.3. The space for place in sociology: the need for a new perspective

Space-time thinking in the social sciences is usually overlooked, whereas space and society are interdependent and strongly connected.

New technologies made possible increasing networking by creating the contraction of space time, which is not the disappearance of place. Places are still there, and they are closer to each-other. This is because space is a social construct and it may, in turn, affect social behaviours. In other words, as Castells (1996) puts it: "Spatial forms and processes are formed by the dynamics of the overall social structure".

For these reasons one important area of focus for social sciences should be the interactions between infrastructure networks, urban spaces and working practices, since "time and space cannot be understood independently of social action" (Harvey 1989: 204).

According to Martinotti (in Nuvolati, 2001: 17): "Sociological analysis has curiously overlooked the topic of mobility and the relations between transport technologies and the social system as a whole". The only exception are immigration studies, which involve another kind mobility and tells little about daily life space-time experience.

The topic has usually been studied from the perspective of long-distance migration flows, whereas it has not from temporary ones' involving daily commuting practices. The emerging idea of flows and network of different actors, when defining the urban, gave new importance to measuring flows instead of stocks: "modern cities are extraordinary agglomeration of flows" (Amin, Thrift 2001: 68).

Sociology needs to address a new perspective from the non-place to the place; whereas economic geography should consider labour and workers as a new privileged subject of investigation.

Indeed cities are like open doors and another way to study the relations among them is to study the flows of people entering and exiting these doors (Perulli 2012). Studying commuting flows, in other words, allow us to find empirical evidence (or not) for the definition of the so-called "macro-region". We advance the claim for a new kind of geographical scale; the geography is something we build continuously through our

[26] Suburbanization and sprawl first took place in Milan, Torino and Genova (Feltrin and Maset in Perulli: 63), whereas starting from 1990s it became consistent along lines parallel to the main communication infrastructures, by creating real urban corridors (Ibidem: 76).

daily practice, so that our purpose is to define a new economic and social geography based on workers' movements and interactions, not just on firms' ones[27].

Since we approach together urban and labour studies, we cannot avoid studying mobility practices. Mobility is definitely a structuring phenomenon we should be aware of from the fact that territories got urbanized along mobility corridors, as we mention in chapter 2 for the Italian case.

In order to better understand the transformation which took place within cities we will refer to the contributions of Martinotti (1993). His papers distinguishes among traditional city and metropolis of 1st, 2nd and 3rd generation. The first traditional object of urban sociology has been the so-called "traditional city" where inhabitants and workers used to coincide. When the transport network expanded urban population was segmented along the daily time line: the diurnal population (working in the city) and the nocturnal one (living in the city). Along with the transition from a producing oriented society towards a consuming-oriented one, a new subject appeared in urban areas: the cityuser, looking for cultural entertainment and amenities. This is the third generation metropolis, mostly crowded by commuters and consumers in the day-time[29].

Despite space and time being the fundamental, material dimensions of human life, academics in social science fields often forget to include them in their analysis.
Inspired by Henry Lefebvre (1958)[28], Allen J. (1999: 56) was one of the first to talk about the "rythms of the city" as the different space-time coordinates forming several urban experiences.

Hägerstrand (1975)[29] along with using the individual human as the unit of study, he also emphasized the importance of time in human activity; "Time has a critical importance when it comes to fitting people and things together for functioning in socio-economic systems," he noted. Hence, a given location may be near an individual, but if a person cannot allocate enough time to travel to it, spatial proximity alone will

[27] The traditional object of economic geography was considered to be the big economic actor, the enterprise. We moved the perspective towards the single worker in order to examine relations among nodes part of the polycentric city-region. [29] According to Barley (2000: 9), every morning (between 7 and 10 a.m.) the number of people in London city increases by 1,3 millions compared to the night time.

[28] The idea of "spatial practice" makes the qualities of urban space embedded in our everyday experience, such as work, probably the most relevant daily experience among all.

[29] Space-time geography has revolutionized the study of transportation accessibility. Throughout the 1980s and 1990s Hägerstrand's model continued to influence fields ranging from city planning to social equity.

not be enough to allow the person to visit it. He elaborated the suggestive idea of space-time path and identified 3 categories of constraints to human mobility: capability, coupling, and authority[30].

A coupling constraint refers to the need to be in one particular place for a given length of time, often in interaction with other people: in other words, your space-time path must temporarily link up with those of certain other people to accomplish a particular task. This kind of constraint is probably the most relevant one for human interactions especially in the work-field and for the specific category of Ipros.

To summarize, the gradual process of tertiarization of the economy and the growing technological development, mainly when referring to capital-intensive jobs, has led to a new definition of time and space in the daily routines. Moreover Ipros often experience daily fragmented times and local[31] mobility flows together with an increasing interprenetation of working and living which makes the boundaries between workplaces and living spaces, working hours and leisure time blurrier (Gotschall and Voss, 2003).

1.3.1 Mobility[32] and infrastructure networks for a new geography of stratification

As has been said repeatedly, both space and time have been transformed under the effect of the information technology: telephony, satellite, cable and mobile have given the biggest contribute to compression of distances, more and more measured in terms of time than space.

Cities and urban regions are important sites articulating the (global and local) movements of people (workers, migrants, refugees, tourists) via complex and multiple systems of physical transportation.

Nevertheless infrastructure networks "are often viewed as <<engineers'-stuff>>, not worth the interest of the social sciences", argues Coutard (1999). On the countrary because they are the assets of the city, they are intrinsic to daily experience of urban life; they shape daily life of people. Transport networks are the material basis for the "space of flow" where "no place exists by itself, since the positions are defined by the exchanges of flows in the network" (Castells 1996: 442). This way the technological infrastructure that builds up the network defines the new economic space where people act.

---

[30] Capability constraints refer to the limitations on human movement due to physical or biological factors. An authority constraint is an area that is controlled by certain people or institutions that set limits on its access to particular individuals or groups.

[31] The term "local" refers to urban-region scale, beyond the single city or metropolitan area.

[32] Urry (2000) suggests the need to reformulate the sociological discipline as a "sociology of mobilities", taking into account the "complex interdependencies between, and social consequences of, these diverse mobilities" (Ibidem :185). [35] This is especially true when we consider that "public and monopolistic models of regulation and ownership are being challenged by waves of privatisation and liberalisation" which generally "means a loss of the redistributive social role implied by such public monopolies" (Graham and Marvin 2001: 102).

Building, setting up and locating an infrastructure it is not a neutral decision.

Today we witness the growing differentiation in the use of time which reflects the tendency towards social polarization and marginalization of some city-users with respect to others.

Castells in 1997 used the term "switched-off territories" when stating that "the planet is being segmented into clearly distinct spaces defined by different time regimes" (Ibidem: 21). Indeed infrastructure networks serve to mediate and construct our diverse experiences of time and space, so that urban populations do actually live at different speeds.

People need infrastructures in order to extend their actions in time and space; the configuration of infrastructure networks is not neutral, but it is involved in sustaining and eventually reproducing "sociotechnical geometries of power" (Massey 1993). If we consider "the growing emphasis on global-local rather than intra-urban connections" (Graham and Marvin 2001: 113), what takes place is the spatial manifestation of the dominant interests[35]. Thus when studying urban life we cannot avoid studying networks and the mobility they support by connecting or disconnecting places[33].

Selectivity is imposed to actors crossing many geographical scales and territories by being daily physically and virtually mobile; technological networks too, despite the rhetoric of universality, are thus always specific and contingent in linking one place with another. They "represent geographies of enablement and constraint" (Law and Bijker 1992: 301).

These systems have traditionally been defined as "public local goods" (Pinch, 1985), freely available to all individuals at equal costs. However today major urban infrastructure networks "are gradually being opened up to private sector participation in the management and provision of services", as Graham and Marvin (2001) denounced.

The liberalisation in the provision of such services is leading to an end the era of national and local monopolies and is paving the way for an increasing spatial segmentation: valued spaces are defined by their fast-track connections (transport, telecommunications) which bypass less favoured places and create "switched-off territories".

To conclude and to remark the importance of the topic with respect to our research, it is good to underline its contribution to at least[34] two different diciplines**.** Economic geography more and more is called to approach a micro perspective (being the worker the unit of analysis) in order to explain economic relations

[33] The worst scenario was predicted by Peck (1996) describing infrastructure development as increasingly supporting new vectors of flow and interaction between highly valued spaces and users locked into highly sophisticated international divisions of labour.

[34] The interplay between infrastructure networks and the diverse mobilities they underpin has often been ignored by different disciplinary approaches (geography, urban studies, sociology, engineering, architecture).

and flows among territories, whereas labour studies should not avoid considering space and time as two main dimensions shaping their object of study. Workers are not abstract entities, they are fully part of the material environment they move in. Labour markets, Welfare regimes, and spatial processes play together. These mechanisms should be understood.

1.4. The role of space shaping working times for Ipros: a case study in Italy

As specified in chapter 1.1.2, the so-called "Ipros" (own-account workers on the tertiary advanced sectors excluding liberal and regulated professions) are the target population of the study[35].

The choice has led towards these workers mainly because of two reasons:

1. They have recently grown in number[36], despite they still represent a small percentage among the total working population.

2. They are the ones experiencing the highest rate of work-related mobility, going beyond short distance commuting routines and framing their life as citizens of the whole "city region".

The rising number of Ipros and the growing pattern towards urbanization on a global scale are two interrelated phenomena, since the knowledge economy finds its natural habitat within the urban environment where professionals usually spend their work-time.

The research is set in Italy, with a specific focus: the north-west Italian region in between Torino, Genova and Milano. These area has been selected because, with respect to national average, it shows the highest rates of self-employment status and of highly skilled people working in advanced service activities. Some evidence is provided by Eurostat LFS aggregated data on regional employment (% of total national rate).

If we consider separately the following criteria such as
% of self-employed[37]
% of workers with tertiary education

[35] We decided to exclude from our analysis professionals enrolled in professional registers and atypical self-employed (para-subordinate contract forms) because our concern is to study the time - space axis in daily working practices and their relevance for housing choices. We only focus on the real own-account workers.

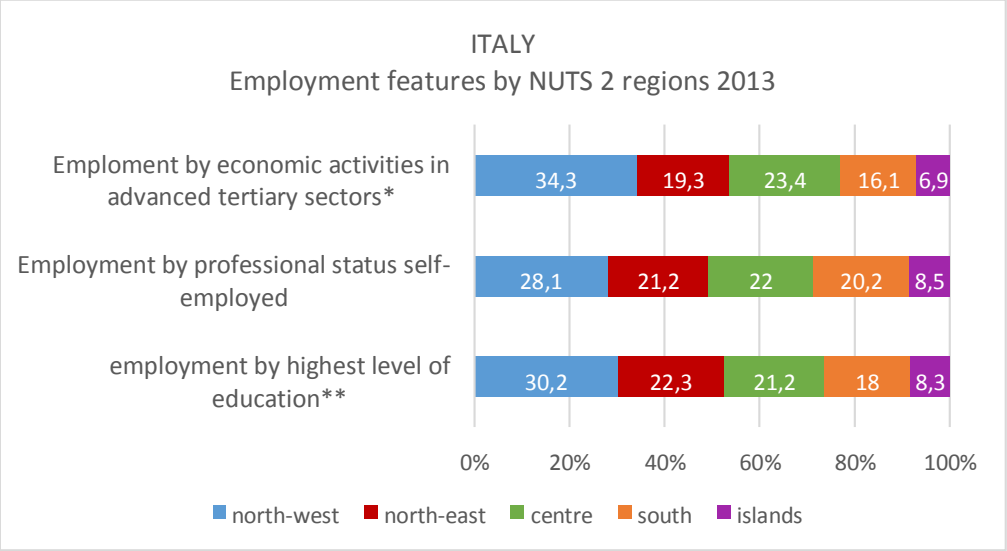
[36] From 2004 to 2013 the EU27 Ipros population grew from fewer than 6.2 million professionals to nearly 9 million, a 45% increase. The growth over that period has been ranging from +12% in Italy up to +93% in Netherlands (Leighton 2013, data from Eurostat LFS).

[37] Aggregated data at regional level do not provide estimates for own-account workers. Anyway the national percentage of own-account workers between 200 and 2011 has grown respectively of +52.8% and + 59.0% in the Netherlands and in Italy (Eurostat, LFS).

% of workers in advanced tertiary sectors

We can easily find out that highest rates for all the 3 criteria are found in the NorthWest area[38] in Italy.

Graph 1. Employment rates in Italy (regional %) by tertiary education, self-employment status and advanced economic activities, 2013



* ICTs + financial and insurance + professional, scientific and technical activities, administrative and support service activities

** Short-cycle tertiary education, bachelor or equivalent, master or equivalent and doctoral or equivalent level (levels ISCED 2011: 5-8)

Source: own elaboration from Eurostat, LFS 2013 (lfst_r_lfemp)

1.4.1 Main research questions and the issue of spatial proximity

The general aim is to describe daily life of these workers. We know very little about them and there is the need for empirical descriptive data. Thus the research project is made up of two different steps. First we would like to provide a static picture of the phenomenon, by mapping the spatial distribution of the whole population of Ipros on the territory, by trying to focus on what we claim to approach the polycentric city region model.

Secondly we would like to track where these workers come from and move to, that is the commuting patterns experienced by a selected sub-sample among the population.

The first point answer the question "Where do Ipros live?", while the second answers to "where do Ipros work?".

[38] More accurately should be noted that the Eurostat classification of NUTS 2 areas includes also Valle d'Aosta among the administrative regions forming the North-West of Italy.

RESEARCH QUESTION 1

This research question cannot be formulated without mentioning the hypothesis of spatial agglomeration and proximity constraint.

Edward Glaeser (1998) highlights the central role of proximity and urban density in enhancing labor concentration, processes of knowledge diffusion, exchange of ideas and job mobility.

Major players in the knowledge economy are the flexible[39] entrepreneurs of knowledge (Sennett 1998) who find their typical habitat in the post-industrial high-speed city and set the proximity rule as governing their daily life in accessing information, and building trust and reputation through personal networks (Duranton 1999).

Social capital matters and it is built locally; it needs physical proximity[40]. Thus economic and social relationships incur spatial transactions costs because many important segments of the economy are dependent on mutual proximity (especially for economic activities and transactions that involve frequent, unpredictable, and constantly shifting face-to-face encounters)[41].

Someone could counter that the development of electronic information and communication systems allows for an increasing disassociation between spatial proximity and the performance of everyday life's functions (among them work)[42]. However, against most diffused forecasts, "the expansion of information and communication technologies has not replaced material flows with immaterial ones" (Martinotti in Nuvolati: 16)[43]. On the contrary urban sprawl enhanced local mobility and short-range up to long-range commuting patterns ("third-generation metropolis"). This is to proof that telecommunication do not automatically lead to dematerialization of social organizations, on the contrary they increase the global speed of society, by increasing the total number of physical moves for people and goods (Martinotti in Nuvolati: 22)[44].

[39] Flexibility is used as synonymous of job temporariness and spatial mobility.

[40] Social interaction could take place *vis à vis* (direct interaction) or be mediated by technology (indirect interaction). Contrasting remote connections with direct ones is misleading: these are complementary behaviours, not competing ones. "However the second type of interaction cannot take place completely avoiding the first one" (Perulli 2012: 30).

[41] Face-to-face communication is a unique type of communication including signals and information otherwise not conveyable (Ekman, Rosenberg 1997). It allows implicit knowledge to be shared and transmitted (Ettlinger 2000).

[42] In the XXI century there has been many claims for the future loss of the local dimension within brand new technological cities (Virilio 1984) and the role of "dis-embedding" produced by new space-time compression.

[43] According to Negroponte (1991) we should consider the other side of the coin: telecommunications does enhance mobility because they allow people to be in touch with their own working places and not to give up their private life.

[44] If we consider time-related-commuting in the United States the average time dedicated to working trips increased by 24% between 1977 and 2009, whereas the average commute trip length increased by 30% in the same period ('Summary of Travel Trends', report from National Household Travel Survey NHTS 2009).

Once more, ICTs development "does not imply the end of the office, but the diversification of working sites for a large fraction of the population, and particularly for its most dynamic, professional segment"[45] (Castells 1996).

Social life is very relevant for Ipros because it is itself a working-place where to establish ties with other workers, get in touch with clients, test trust and reliability and built a reputation (Boden in Giddens 1990). Restaurants, club, sports club, sponsored events, official meeting are the places where social-working life takes place and weak ties may establish (Granovetter, 1973)[46].

This is true especially if we consider that most of commuting is determined by consuming patterns or provision of business services[47]; these require face-to-face performances and they may find their natural environment in informal settings during extra-working times.

To summarize:

- Social capital and mutual trust are built locally and require extensive face-to-face contacts
- The access to information mostly depends on physical interaction
- «High-skilled workers are more productive when clustered together» (Storper 2013)
- The increasing interprenetation of working and living makes the boundaries between workplaces and living spaces, working hours and leisure time blurrier (Gotschall and Voss, 2003)

For these reasons, assuming the importance of spatial proximity, our hypothesis is that:

I. Ipros tend to live close to each-other

II. Ipros prefer living in dense urban areas with specific features

Our first research question is meant to test it by

I. detecting specific clusters of Ipros living within the urban region

II. identifying specific localities or "central areas" where the clusters should form

[45] It is noteworthy that Castells (1996) argues that, especially in the case of information technology activities, spatial proximity is a necessary material condition for the existence of 'milieu of innovation' because of the interactive nature of the innovation process. "Once established, *milieux* of innovation both compete and cooperate between different regions, creating a network of interaction that brings them together in a com- mon industrial structure beyond their geographical discontinuity."

[46] As Mark Ganovetter describes it, weak ties involve less frequent interaction, lower levels of emotional intensity and intimacy, and are less homophilous. Consequently, weak ties serve an important "bridging" function among different social groups/ networks, opening paths for the rapid and efficient exchange of opportunities and information across social distance.

[47] Face-to-face contacts are required by self-entrepreneur in tertiary sectors both when working in (providing specialized services to businesses) and when accessing services for themselves.

According to our assumption self-employed people should prefer living in central areas, core economic centres; the academic literature provide some suggestions for its definition. Some relevant factors enhancing spatial convergence of economic actors are access to richer physical infrastructures and improvement in information flows through local labor markets where huge quantities of information are informally daily created and circulate and new social encounters and experiences endlessly occur.

For all these reasons most entrepreneurial efforts do locate within big urban centres.

Therefore what we define "central areas" are urban nodes where "Position and centrality are configured less by geographical location with respect to 'downtown' than by conditions of building and places with respect to global-local networked infrastructures like international airports, high speed rail and port terminals, optic fibre links, broadband internet 'pipes' and satellite terminals" (Graham and Marvin 2001:122). The technology of infrastructure have increasingly become the organisation principle to everyday life and they are part of the locationally-dependent constraints or costs affecting housing choice.

Ipros provide advanced services and at the same time ask for them. Thus our idea of central area is very close to the idea of hub-city (Perulli 2012: 31, 40) which should be accessible and open, providing a set of good material and immaterial transport networks[48], hosting major research centres and firms, being open to international flows (airports, harbours, immaterial networks etc.). In other words they are the nodes of the network identified by Castells (1996:443) as requiring an "adequate technological infrastructure, a system of ancillary firms providing the support services, a specialized labor market, and the system of services required by the professional labor force". The ideal place for knowledge workers to locate themselves and where to find the greatest opportunities for personal enhancement, social status, and individual self-gratification.

RESEARCH QUESTION 2

Its goal is an explorative and descriptive one. It addresses the main issue of spatial mobility (we deeply discussed in chapter 1.3).

Our main purpose is to be able to track work-related mobility pattern for a sample of Ipros, in terms of weekly frequencies and commuting distances.

How are these patterns influenced by location choice made by Ipros (research question 1.II)?

[48] "The geographical proximity within cities is no guarantee of meaningful relations or connections" (Graham and Marvin 2001: 204); time-space compression created by infrastructures can be very uneven and partial distinguishing between central areas and marginal ones.

The underlying hypothesis is that socio-demographic variables such as age and gender are very related to the amount of disposable income and social capital that these workers can benefit from.

We expect to find, with respect to commuting variables, an internal segmentation among Ipros in terms of age and gender. Despite the use of a general wide term ("Ipros"), this is not a very homogeneous category, as we have noted from Italian data in chapter 1.1.2. We believe it is worth to focus deeper on this cleavages in order to challenge those studies denying the existence of this growing very complex universe within the labour market stratification. Who are the individuals we are going to study? Despite they work in new economic leading sectors, they are not part of the new economic elite; they could experience, even in their spatial distribution, forms of marginalization and partial exclusion[49].

Another added value from this description would be very strongly related to the confirmation of the hypothesis that a real urban region does exist in the North-west of Italy and its existence could actually be detected when measuring long-distance internal flows of people[50].

1.4.2 Methodology and the lack of data

Little has changed compared to when Sassen (2000: 144) wrote: "current datasets are quite inadequate for addressing this issues at the level of the city". Most of today surveys address only places of residence, providing the image of the "sleeping city" (Martinotti in Nuvolati:16), while data on mobility (from origin to destination and vice versa) are still scarce and lack socio- demographic information. Official statistics is still anchored to the administrative units of the territory[51], despite the urbanity is no more to be conceived as a social isolated object, but rather as an interacting node in a wider network on a macro-urban scale.

When it comes to the research design, then, the lack of available secondary sources is something we have to deal with.

In order to map the spatial distribution of Ipros (research question 1), we have to rely on responses provided by a first-hand survey data collected on a sample of Ipros, since no available census data are there.

---

[49] Wacquant (1996) first introduced the idea of urban "advanced marginality" created by advanced capitalism, fragmentation and desocialization of labor, casualization of employment.

[50] This would actually give some evidence in order to refuse the definition provided by OECD and the European Commission (2011) of "The Larger Urban Zone" as "the city and its commuting zone". Indeed in this case the long distance interregional commuting patterns are not detected.

[51] Available Italian official data sources (and existing surveys) are largely inadequate to study the growing territorial mobility of urban population (especially commuting workers and "city-users") beyond official administrative borders. Census data from 2011 are the only available source, having many limitations.

The second research question, concerning mobility paths, involves as well analysis of questionnaire data[52] concerning the distance and the places, work-related movements and not[53].

To contact the required sample we will mainly get support from ACTA ('Associazione Consulenti Terziario Avanzato'), a national association founded in 2004 in order to represent the interests of Ipros workers in Italy. The association is part of the international network EFIP ('European Forum of Indipendent Professionals').

At a later time, in order to be able to give proper interpretation to our results we will compare locational choices and mobility patterns with macro territorial data related to available infrastructures, density of businesses and innovation-oriented market. The mapping might give evidence of existing territorial inequalities by testing if "infrastructure networks more intensively and actively connect valued places, while at the same time progressively withdrawing and disengaging from less valued places" (Graham, Marvin 2001: 138).

[52] Despite the retrospective survey could not provide very reliable data, the other possible detecting technique would be using Gps tracking tools; however this is a very demanding one in terms of costs and times (to build up the available technology apparatus and to find volounteers to monitor). Last option would have been launching e-mail diary entries, which was not chosen since it is proved by literature to have low response rates.

[53] Pilot interviews can be useful to define the best space and time unit describing mobility phenomenon for Ipros. Another useful reference would be the 'Study of Mobility in the Netherlands' (OViN), a national Dutch annual survey collecting data from 185 onward about people mobility; among most relevant variables we have Kilometres travelled by age, gender, occupation, income, mode of transport, province of departure, month, weekday, time of departure, motive.

References

Allen J., et al. (edited by) (1999), *Unsettling cities*, London, Routledge

Amin A., Thrift N. (2001), *Cities. Reimagining the urban*, Cambridge, Polity Press

Barley N. (2000), 'People', in Barley N. (edited by), *Breathing cities*, Basel, Birkhäuser

Bailey N. e Turok I. (2001) 'Central Scotland as a polycentric urban region: useful planning concept or chimera?', *Urban Studies ,* 38(4) : 697-715.

Bell D. (1973), *The Coming of Post - Industrial Society*, New York, Basic Books

Bertolino S., Cominu S., (2011) , *Vecchie e nuove professioni a Milano: monadi, corporazioni o terzo stato in cerca di rappresentanza?,* Milan, AAster Consortium research report

Bijker W., Law J. (1992), *Shaping Technology / Building Society: Studies in Sociotechnical Change*, , Cambridge, MIT Press

Boden D. (1990), 'Papers on trust', in Giddens A., *The Consequences of Modernity*, Stanford, Stanford University Press

Bögenhold D., and StaberU. (1991), 'The Decline and Rise of Self-employment ', *Work, Employment and Society,* 5(2): 223-239.

Bologna S., Fumagalli A. (edited by) (1997), *Il Lavoro Autonomo di Seconda Generazione. Scenari del postfordismo in Italia*, Milano, Feltrinelli

Bologna S., Banfi D. (2011), Vite da freelance. I lavoratori della conoscenza e il loro futuro, Milano, Feltrinelli

Bonomi A., Abruzzese A. (2004), *La città infinita*, Milano, Bruno Mondadori

Brenner N. (1999), 'Globalisation as reterritorialisation: the re-scaling of urban Governance in the European Union', *Urban Studies* , 36(3): 441-451.

Brenner N. (1998), 'Global cities, glocal states: global city formation and state territorial restructuring in contemporary Europe, Review of International Political Economy 5 (1): 1-37

Brenner N. (2010), *The <<cityness>> of the city: what is the appropriate unit of analysis for comparative urban studies?*, Lugano, Globalization of Urbanity, USI-Lugano

Brenner N. (edited by) (2014), Implosion / Explosions. Towards a study of planetary urbanization, Berlin, Jovis

Bryson J. R., et al. (1997), 'The creation of small business service firms in post-industrial Britain', *Small Business Economics, 9*: 345-360

Castells M. (1996),*The information Age: Economy, Society and Culture*, Vol. I, Oxford, Blackwell

Colleoni M. e Zajczyk F. (2004), 'La mobilità nello spazio e nel tempo delle popolazioni urbane. Una lettura sociologica', in *Sociologia Urbana e Rurale*, 72;

Coutard O. (1999), The Governance of Large Technical Systems, New York, Routledge

Drucker P. F., (1993), Post-capitalist society, New York, Routledge

Duranton G. (1999), 'Distance, Land, and Proximity. Economic Analysis and the Evolution of Cities', in *Environment and Planning A*, 21 (1): 2169-2188

Ekman P., Rosenberg E. (1997), *What the face reveals*, Oxford, Oxford University Press

Esping-Andersen G., Regini M. (2000), *Why De-Regulate Labour Markets?,* Oxford*,* Oxford University Press

Ettlinger N. (2000), *Frontiers of Flexibility and the Importance of Place and Space*, Paper, Department of Geography, Ohio State University

Friedmann J. (1995), 'Where We Stand: a Decade of World City Research', in Knox P.L. and Taylor P.J. (eds), *World Cities in a World-System* , Cambridge, Cambridge University Press

Garavaglia L., Pennati D. (2014), 'Città medie e sistemi di corridoio', paper presented at Fall School in Local Development , Asti 27-28 November 2014

Giddens A. (2006), *Europe in the Global Age*, Cambridge, The Polity Press

Glaeser E. (1998), 'Are cities dying?', *Journal of Economic Perspectives 12*: 139-160

Graham S., Marvin S. (2001), Splintering urbanism: Networked Infrastructures, Technological Mobilities and the Urban Condition, London & New York, Routledge

Granovetter, Mark S. (1973), 'The Strength of Weak Ties', *American Journal of Sociology 78* (6): 1360-1380.

Hägerstrand T. (1975), 'Space, time and human conditions' In A. Karlqvist, et al. (eds.) , *Dynamic Allocation of Urban Space*, Lexington, Saxon House Lexington Books.

Hall P., Pain K. (2006), *The Polycentric Metropolis. Learning from mega-city regions in Europe*, London, Earthscan

Harvey, D. (2008) 'The right to the city', *New Left Review*, 53: 23–40.

Harvey D. (1989), *The Condition of Postmodernity. An Enquiry into the Origins of Cultural Change*, Oxford, Wiley-Blackwell

Herrschel T. e Newman P. (2002), *Governance of Europe's city regions: planning, policy and politics*, London, Routledge

Kloosterman R.C, and Musterd S.,2001. and S. Musterd, 'The Polycentric Urban Region: Towards a Research Agenda', Urban Studies, 38(4): 623-633

Krugman P. (1997), *Pop Internationalism*, Cambridge, MIT Press

Le Febvre H. (1970), *La révolution urbaine*, Paris, Gallimard ed.

Le Febvre H. (1974), La production de l'espace, Paris, Anthropos

Le Galès P. (2002), *European Cities. Social Conflicts and Governance*, Oxford, Oxford University Press

Leighton P. (2013), The Rise of Europe's Independent Professionals (iPros), EFIP - PCG

Luber S., Leicht R. (2000), 'Growing self-employment in Western Europe: An effect of modernization?', *International Review of Sociology,* 10 (1): 101-123

Marra E., Mela A., Zajczyk F. (2004), 'Tempi difficili per la città', in *Anni in salita. Tempi difficili per gli italiani,* Amendola G. (a cura di), Milano, Franco Angeli

Martinotti G. (1993), *Metropoli*, Bologna, Il Mulino

Massey D. (1993), 'Power-geometry and a complessive sense of place', in Bird J, et al. (edited by), Mapping the futures: Local Cultures and Global change, London, Routledge

Meijers E. (2007), Sinergy in Policentric Urban Regions. Complementarity, organizing capacity and critical mass, Alblasserdam, Haveka

Negroponte N. P. (1991), 'Product and services for computer networks', *Scientific American*, 265 (3)

Nuvolati G. (2002), *Popolazioni in movimento, città in trasformazione. Abitanti, pendolari, city-users, uomini d'affari e flâneurs*, Bologna, Il Mulino

Perulli P. (edited by) (2012), *Nord. Una città-regione globale*, Bologna, Il Mulino Studi e Ricerche

Perulli P., Pichierri A. (edited by) (2010), *La crisi italiana nel mondo globale. Economia e società del Nord*, Torino, Einaudi

Pichierri A. (2002), *La regolazione dei sistemi locali. Attori, strategie, strutture,* Bologna, Il Mulino

Piore M., Sabel C. (1984), *The Second Industrial Divide*, New York, Basic Books

Porter M. (1995), 'The Competitive advantages of the inner city', *Harvard Business Review*, May-June 1995: 53-71

Ranci C. (edited by) (2012), *Partite Iva. Il lavoro autonomo nella crisi italiana*, Bologna, Il Mulino

Rapelli S. (2011), *European Ipros - A Study*, PCG

Reyneri E. (2011), *Sociologia del mercato del lavoro*, vol II. *Le forme dell'occupazione,* Bologna, Il Mulino

Rodgers G. (1989) 'Precarious Work in Western Europe' in G. Rodgers and J.Rodgers (eds), *Precarious Jobs in Labour Market Regulation: The Growth of Atypical Employment in Western Europe*, Belgium, International Institute for Labour Studies: 1-16.

Sassen (1994), *Le città nell'economia globale*, Il Mulino, Bologna.

Sassen S. (2001), *The Global City: New York, London, Tokyo*, Princeton University Press

Scott A. j. (edited by) (2001), *Global City-Regions: Trends, Theory, Policy*, New York, Oxford University Press

Scott A. j. (2011), 'A world in emergence: notes toward a resynthesis of urban-economic geography for the 21st century', *Urban Geography*

Scott A.J., Storper M., (2003), 'Regions, globalization, development', *Regional Studies*, 37: 579-593

Sennett R. (1998), The corrosion of character. The personal consequences of work in the new capitalism, New York, Norton

Simpson W., van der Veen A. (1992), 'The economics of commuting and the urban labour market', *Journal of Economic Surveys*, 6 (1): 45-62

Soja E. (1992), 'Inside Exopolis: Scenes from Orange Country' in *Variations on a Theme Park,* Sorkin M. (edited by), N.Y, Noonday Press

Urry J. (2000), *Sociology beyond societies: mobilities for the twenty-first century*, London, International Library of Sociology, Routledge

Van der Burg A. J., Vink B. L. (2008), '*Ranstad Holland 2040*', paper for the 44th ISOCARP Congress Virilio P. (1998), *Lo spazio critico*, Bari, Dedalo ed.

Loïc J. D. Wacquant (1996), 'The Rise of Advanced Marginality: Notes on Its Nature and Implications', *Acta Sociologica, 39*, (2): 121-139

Chapter 2

Survey design and data collection

2.1. The need for an explorative survey

Our research design is mainly affected by the lack of available secondary sources: if we rely on official statistics we know very little about Ipros and freelancers, this is why if we want to study this (not yet well depicted) population there is no other way than collecting first hand data by implementing a specific survey.

Our general goal is mostly a descriptive one and we investigate 6 key dimensions such us working activity, residential condition, socio-economic status, use of ICTs, work-related mobility, life style and time-use patterns.

Some of these dimensions are part of very specific existing official national surveys such as surveys on mobility or labour force[1]; therefore even when referring to the same country these data are collected separately on different samples and are not comparable. Furthermore their general target is the national population and they hardly allow the focus on specific subgroups such as freelancers (i.e. Istat census on Mobility).

Since our aim is to collect comparable data on such dimensions which mainly shape Ipros' life, we need to design and fulfil our own survey.

After collecting micro data we will introduce in the analysis macro territorial data (land use) with reference to accessible housing market, net of available transport infrastructures and the density of businesses location. Being able to locate each respondent on the map will allow to detect for phenomena of spatial segregation within areas which are isolated from the economy and labour market of the city region (by having lower levels of transport and communication connections, universities and firms linkages, research centres, firms' clusters and so on).

[1] i.e "Study of Mobility in the Netherlands" (OViN), Census data on mobility in Italy, Labor Force Survey (LFS) in Italy and in the Netherlands.

2. 2. Designing the survey

There exist different strategies and methods for collecting survey data.

Choosing one method with respect to another usually depends on research objectives and resources limits (more frequently related to available time and money).

More specifically, when choosing we cannot avoid addressing the following issues:

- goals of the survey and types of information needed
- target population and sample
- questions wording
- mode of administration (face-to-face, telephone, e-mail, web)
- data editing and coding.

2.2.1 Target population

Our target population are the so-called Ipros, self-employed workers in advanced services[2].
In order to be better able to have another term of comparison, the survey is designed to provide a "quasi control sample"[3] made up of professionals, whose formal recognition is granted by their enrollment in professional associations. This is the main difference distinguishing Independent Professionals and Professionals; the first being part of the new knowledge economy[4], whereas the second being involved with much more traditional professions already recognized in long-lasting economic sectors (such as architects, journalists, labor consultants, lawyers and so on).
Our main hypothesis is to be able to detect for at least two cleavages shaping internal segmentation within this heterogeneous group: age (and stage of life) and the belonging (or not) to professional association. To make it easier the question is whether working conditions (timing,

[2] We have already discussed our definition; this is the case when "it is easy to define the population in conceptual terms but difficult to do so in operational terms" (Blair et al. 2014:109)

[3] The control group is the main element of experimental designs, which is not our case. However it is useful to assess our research questions on two comparable samples which are meant to be very similar and homogeneous except from the affiliation to professional associations.

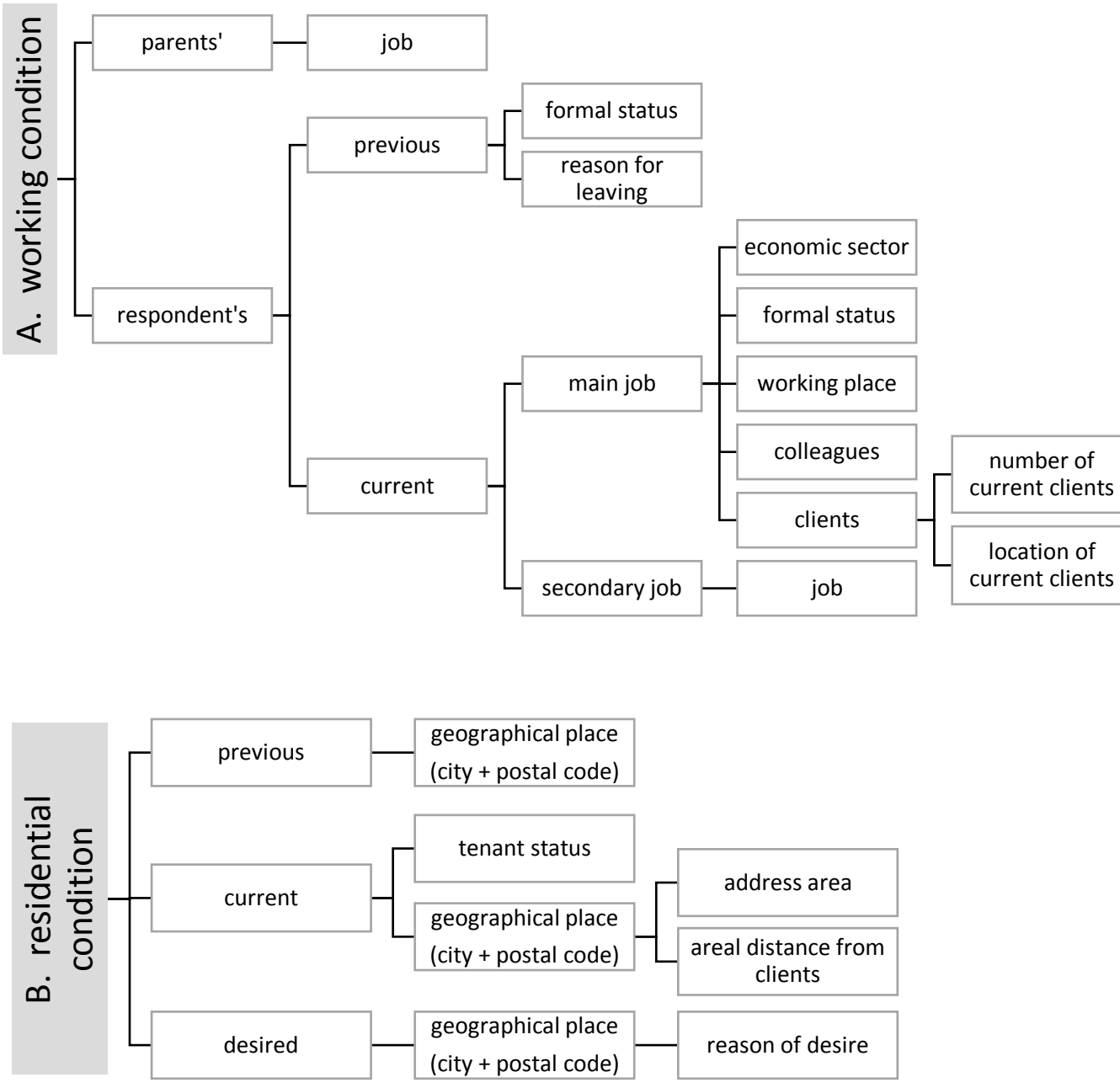
[4] According to the World Bank, knowledge economies are defined by four pillars: institutional structures that provide incentives for entrepreneurship and the use of knowledge, skilled labour availability and good education systems, ICT infrastructure and access, and an innovative landscape that includes academia, the private sector and civil society. The European Commission has addressed these different goals both in Lisbon Strategy 2000-2010 and the most recent Europe 2020.

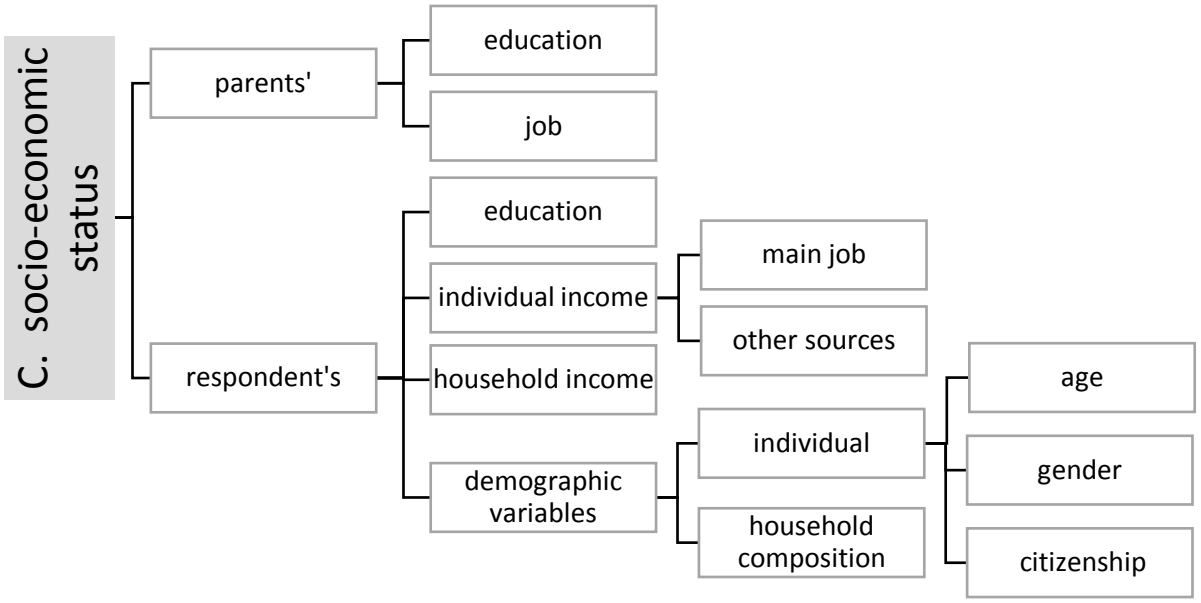
work-related mobility patterns, number of clients, income) and residential patterns might be similar depending on stage of life or the belonging to a professional association.

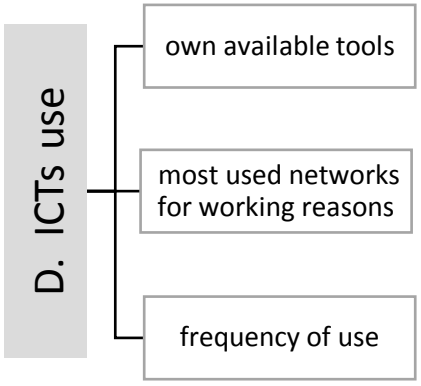
2.2.2 Contents and main dimensions

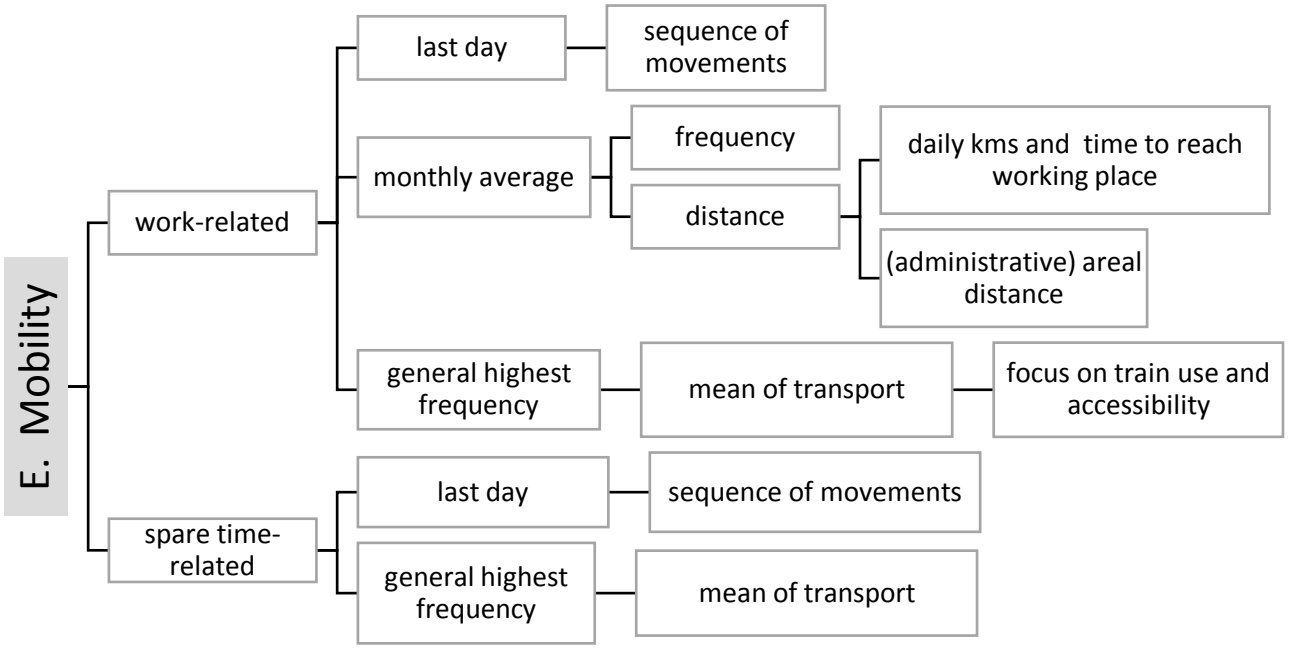
The general goal of our survey has been defined as an explorative one, investigating 6 dimensions of Ipros' life. Namely working condition, socio-economic status, residential condition, ICTs use, mobility and lifestyle.

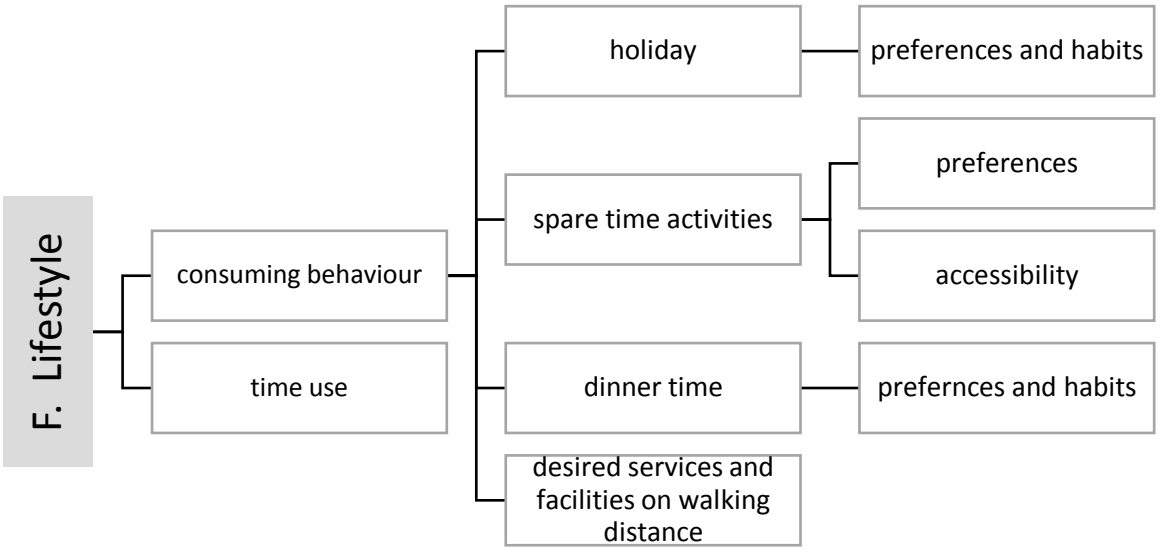
Graph 1. Tree diagram of the questionnaire: key dimensions from A to F and sub-dimensions




C. socio-economic status
parents'
education
job
respondent's
education
individual income
main job
other sources
household income
demographic variables
individual
age
gender
household composition
citizenship


D. ICTs use
own available tools
most used networks for working reasons
frequency of use


E. Mobility
work-related
last day
sequence of movements
monthly average
frequency
distance
daily kms and time to reach working place
(administrative) areal distance
general highest frequency
mean of transport
focus on train use and accessibility
spare time-related
last day
sequence of movements
general highest frequency
mean of transport



2.2.3 Research goals and techniques of data analysis

Our research questions involve studying the locational patterns of Ipros (dimension B) and their work-related mobility and travel behaviour (dimension E)

In order to better answer these questions we will apply respectively two different techniques of analysis.

The first one (chapter 3) considers socio-economic variables in order to identify internal segmentation. We claim that there are mainly two cleavages shaping internal segmentation within this heterogeneous group: age (and stage of life) and gender.

Furthermore we focus on mobility patterns and travel behaviours in relation to residential location and try to eventually identify phenomena of spatial exclusion. Gentrification has been seen as the spatial expression of class inequality so far; our argument is that this could be true no more. Who are the gentrifiers today? Are Professional workers a real homogeneous class? This could not be the answer since they are a very diverse group and there are some of them holding economic "marginal" positions which could in turn being reflected in patterns of spatial exclusion.

These topics are explored by using descriptive statistics and simple regressions tecniques.

The second one (chapter 4) belongs to the wide family of the so-called "**spatial statistics**"[5] and focus on the analysis of areal data which allow to define measures of agglomeration, clustering and spatial exclusion.

In this situation the variable of interest has values within a fixed set of areas (such as districts, census zones etc.) covering the study area and we consider the issue of spatial autocorrelation so that observations are non-independent over space. In other words "data values are observations associated with a fixed number of areal units" (Fischer, Wang 2011: 5) so that we can describe spatial distribution of phenomena, discover patterns of spatial clustering and eventually identify outliers.

This is crucial to us, since we can actually test the hypothesis of spatial proximity and identify "central areas" where, according to the theory[6], clusters should form. Another way to look at this process is to talk about "gentrification", a process "deeply rooted in social dynamics and economic trends" (van Weesep 1994: 80). According to a wide stream in the post-industrial literature main actor of this process should be the middle class (Hamnett 1991, Smith 2002), often view as the new "creative class" (Florida 2003, 2005) .

The term gentrification was first coined by Ruth Glass (1964) with specific reference to class transformation of urban space, nevertheless the labour market, the occupational structure has changed and together gentrification itself has mutated form: "it is the loss of manufacturing employment and the increase in the service employment which led to an expansion in the amount of professionals with a disposition towards central-city living" (Lees et al. 2008 :90). Davide Ley (1972) and, later on, Chris Hamnett (1991) analysed gentrification process in the context of the new postindustrial city marked by "deeper changes in the structure of production, the changing division of labour, and the rise of a locationally concentrated service class[7]" (Hamnett 1991: 177). Hamnett's thesis is also known as the "professionalization thesis" as he argued for "the increase proportion of professional workers in a number of large cities with a strong financial/producer service base" (Hamnett 1994: 407).

[5] Spatially-referenced data is becoming increasingly important in the social sciences with a specific focus on statistics, as "GIS analysis rarely show estimates of confidence or other indicators of the effects of data quality" (Sweeney, Konty 2004: 2). In broad terms we can define spatial analysis as the quantitative analysis of spatial phenomena that are located in geographical space (Bailey and Gatrell 1995).

[6] At the macro level there are trade-offs between space and accessibility that structure different residential patterns: "In the urban environment the attractiveness of land is based mainly on location, accessibility, and the labour and technology devoted to improving a site (…) centrality and accessibility are valued" (Lees et al. 2008: 51).

[7] Mainly due to the desire of minimizing commuting times and the greater ability to afford cultural and social attraction of life in the central and inner city (Hamnett 2003).

"Local clusters of transnational corporate services and headquarters generate trasnational circuits of labor migration amongst itinerant professionals and freelance employment contracts" (Lees et al. 2008: 80).

The crucial point is that professionals provide advanced services but at the same time ask for them, so that according to both the production and from the consumption paradigm they are supposed to settle where these advanced services are provided.

2.3 Collecting data

Evaluating the feasibility of an online, mail, face-to-face or telephone survey mainly depends on resource factors. In this specific case time and money (especially for hiring interviewers) constitute the biggest limit to our research, so that internet survey proved to be the fastest and least expensive way of administering the questionnaire. Another reason for implementing web suveys is that they "are good for highly educated respondent groups" (Blair et al. 2014: 74) with high level of Internet access; both these two characteristics are peculiar of the population of Ipros and professionals.

However this method has severe sampling limitations and usually shows lower response rates than other modes; for this reason we will implement a combination of methods to enlarge the sample number.

2.3.1 Sampling strategies and specific issues for web surveys

The boundaries of target population are ideally defined by the following characteristics:

- own account workers (IPros and Professionals)
- working in the advanced service sector
- with high level of education

In our case study, it is not easy to measure and quantify with high degree of certainty the population of interest according to these over-mentioned characteristics. Estimating the size of the population is the harder task since it does not appear clearly from available official data.

Hitherto the only chance to provide a rough estimate of the proportions of population's demographic characteristics is to draw on official numbers provided by Istat in the national Labour Force Survey (2011).

We first filtered the whole sample by attributes[8] indicating

- independent working time (C6=1)
- independent working place (C7=1)
- being a collaborator or self-employed (DIPAUT=2 | DIPAUT =3)
- being a professional or own-account worker (POSPRO=8|POSPRO=9)

Among this wide subsample (2.193.511 cases) we extracted a sample including only those people above 25 years old, with minimum secondary education and working as specialized professionals or high level technicians[9].

The final numerosity of our subsample is 863.812 cases. The table below shows absolute numbers and cell percentages (the percentage is calculated on the total number of cases).

Table 1. High specialized professionals and higher level technicians by gender, age and educational level

				Highly specialized professionals		*Higher level technicians*		*total*	
			Age	*a.n*	*% by cell*	*a.n*	*% by cell*	*a.n*	*% by cell*
	Male	*Tertiary education*	*25-34*	37.654	4,3	10.157	1,2	47.811	5,5
			35-44	90.920	10,5	17.124	2	108.044	12,5
			45- 54	59.000	6,8	6.630	0,8	65.630	7,6
			55-64	47.801	5,5	5.621	0,6	53.422	6,1
			65-74	17.236	2	1.422	0,1	18.658	2,1
			75 +	5.106	0,5	629	~ 0	5.735	0,5
		tot		257.717	**30**	41.583	4,8	299.300	**34,6**
		Secondary education	*25-34*	5.313	0,6	48.409	5,6	53.722	6,2
			35-44	7.780	0,9	87.113	10	94.893	10,9
			45- 54	12.775	1,5	78.072	9	90.847	10,5
			55-64	6.846	0,8	41.712	4,8	48.558	5,6
			65-74	1.448	0,1	9.878	1,1	11.326	1,2
			75 +	626	~ 0	1.032	0,1	1.658	0,1
		tot		34.788	**4**	266.216	**30,8**	301.004	**34,8**

[8] Names and codes of the filter variables are C6, C7, DIPAUT, POSPRO.

[9] As already remarked there are good reasons for including specialized technicians in the calculations (see previous chapter)

Tot	600.304	600.304		292.505	**33,8**	307.799	**35,6**	600.304	**69,5**
	Female	*Tertiary education*	*25-34*	34.353	4	8.363	0,9	42.716	4,9
			35-44	78.006	0,9	7.210	0,8	85.216	1,7
			45- 54	31.697	3,6	4.404	0,5	36.101	4,1
			55-64	9.240	1	747	~ 0	9.987	1
			65-74	3.739	0,4	225	~ 0	3.964	0,4
			75 +	346	~ 0	-		346	0
		tot		157.381	**18,2**	20.949	**2,4**	178.330	**20,6**
		Secondary education	*25-34*	3.607	0,4	14.613	1,7	18.220	2,1
			35-44	5.447	0,6	20.550	2,4	25.997	3
			45- 54	6.567	0,7	20.676	2,4	27.243	3,1
			55-64	3.453	0,4	7.809	0,9	11.262	1,3
			65-74	327	~ 0	1.867	0,2	2.194	0,2
			75 +	262	~ 0	-		262	0
		tot		19.663	**2,2**	65.515	**7,6**	85.178	**9,8**
Tot	263.508	263.508		177.044	**20,5**	86.464	**10**	263.508	**30,5**
Total	863.812	863.812		469.549	**54,3**	394.263	**45,7**	**863.812**	100

Source: Istat, LFS 2011

Designing a representative sample based on these figures is fundamental for the idea of statistical inference. The notion of representativeness is essential when you want your sample to provide an accurate representation of the broader population.

According to the typology drawn by Couper (2000) web surveys may involve different types of nonprobability and probability samples.

The key to be able to draw probability samples is to have access to a frame, that is a list of the members of the population. For web or mail survey the list should include email addresses to send invitation to answer the survey, but it is also possible that the traditional recruiting method is applied (by sampling from a given list or applying random-digit dialing) to select a general population sample and those without internet access are provided with access (in order to avoid coverage error).

Unfortunately with respect to our research design the only suitable way to have a quasi-probability sample is the so-called “intercept survey” (Tourangeau et al. 2013: 12). In this very specific case “a random or systematic sample of visitors to a particular Web site is selected over

a specified time frame and invited to complete the survey". The population frame consists of the visitors[10] to the Web site during the specified time window, despite the sample is defined as "random" as we cannot design our sampling strategy in advance.

Although the coverage error is overcome, the restricted population does not allow to draw conclusions to a broader population. We need to enlarge it.

How to avoid, then, the statistical consequences of non-probability sampling (biased estimates of the corresponding population proportions[11])? Complete omission (that is zero probability of inclusion) of some portion of the population of interest is generally avoided since Ipros are supposed to have large access to internet and the web. Among those with nonzero probability of inclusion, however, we may find differences in the inclusion probabilities which covary with the survey variable of interest.

As there is non guarantee that the volunteer members resemble the larger population of Ipros on the survey variables[12], we have to rely on post-stratification adjustment to reduce selection bias and make the sample more similar to the population[13].

The most common method for adjustment is known as "post-stratification" or "cell weighting"; "this procedure adjusts the sample weights so that the sample totals conform to the population totals on a cell-by-cell basis" (Kalton and Flores-Cervantes 2003). A subsample of respondents is selected from a particular survey and so that population proportions are correctly reproduced in the sample.

When figures are not available for every adjustment cell formed by crossing the auxiliary variables or there may be very few participants in a given cell, raking might be preferred to post-stratification. This second method adjust the sample to the marginal totals for the auxiliary variables, not to the cell totals (Tourangeau et al. 2013 : 26).

---

[10] Actually this sort of sample design provides the probability sample of the population of visits to the site rather than the population of visitors, but "cookies can be used to prevent multiple selections of the same visitor" (Couper, 2000).

[11] The size and direction of the bias depends on the proportion of the population of interest with no chance of inclusion in the sample and differences in the inclusion probabilities among the different members of the sample (subgroups). In mathematical terms: Bias= $E(\bar{y} - \bar{Y})$ and Bias = $P_0(\bar{Y}_1 - \bar{Y}_0) + \frac{Cov\ (P,Y)}{\bar{P}}$

[12] Variables shown in table 1 (age, gender, education, profession).

[13] In this very specific case we should consider that we cannot rely on census data. Our reference population is itself a sample officially estimated by the national Labor Force Survey. This certainly introduces another source of error in the estimates of the percentages cell by cell (post-adjustment) but it is the only available data we have so far.

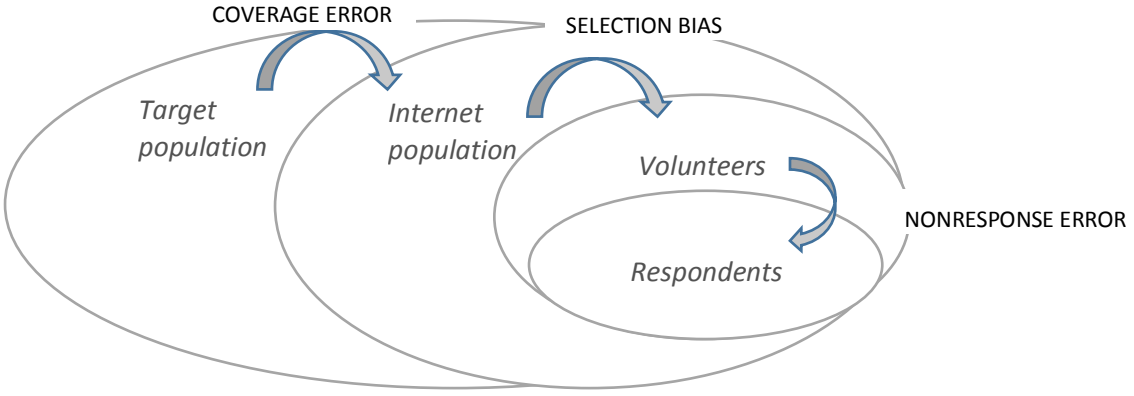
Table 1 shows clearly that numerosity by cell might be problematic in several cases (for elder people and females percentage is often below 1), whereas considering marginal totals for one or more crossed variables (es gender, gender-profession, gender-education) results in a broader number of cases.

2.3.2 Possible sources of error

In order to better classify the errors, it is useful to adopt the "total survey error" framework first specified by Groves (1989). According to this perspective errors are first divided in two categories: those errors affected by sampling coverage (how well the respondents represent the population?) and those affected by the validity and accuracy of individual answers (how well the respondents answer the questions?).

In other words we have errors caused by non observations as well as errors caused by observations. The former occur when we rely on volounteer web panels that may not constitute a representative sample of the population of interest (coverage error, selection bias and nonresponse error[14]). It is pretty clear that this will lead to biased survey estimates if respondents differ from their unobserved counterparts on the attributes being measured in the survey. "If the likelihood of recruitment (and selection) is related to values on the variable of interest, this will introduce error into the estimates"; a typical example being self-selection bias (Ibidem: 24).

Graph 2. Possible source of error from target population to respondents.



Source: Tourangeau, Conrad, Couper (2013: 98)

[14] Also known as "nonsampling error" (Blair et al. 2014: 89) they are caused by units non response: web survey are easy to decline and so achieve lower response rates than interviewer-administered surveys. For this reason we introduced a prepaid incentive for survey participation that motivate respondents; "both theory and observation confirm the importance of incentives (including but not limited to monetary incentives) for participation in surveys" in all modes, including the web (Singer, Ye 2013).

Anyway if the relationship between the self-selected volounteers' characteristics and those of the general population is not known nonresponse cannot really be tied to nonresponse error (Ibidem: 4). In this case the lack of a probability sampling and the common low participation rates do not allow any generalization beyond the participants.

On the other hand validity and **reliability** are typical examples of **measurement errors**, defined as "the difference what respondents report when they answer a survey question and the true value of the attribute being measured" (Ibidem: 3). Examples of this kind of errors are due for example to primacy effect or social desirability.

In mathematical terms we would say that each observation (Y_{iA}) collected from respondent I under method A is equal to its true score ($\mu_{i)}$ plus an error under the mode of data collection A (ε_{iA}):

$$Y_{Ia} = \mu_i + \varepsilon_{iA}$$

Deviations from the true score may rise from a number of sources, including question wording, order, format etc. (Ibidem: 58).

In order to prevent these kind of errors we designed our survey by

- including a certain degree of interactivity which allows the survey to react to the answers of the respondents (conditional routing);
- prompting for missing data;
- including definitions and clarifications;
- presenting a single group of related questions on each page (instead of a single scrolling Web page);
- providing different text formats for instructions, questions, clarifications.

References

Bailey, Gatrell (1995) Interactive Spatial Data Analysis, Longman,

Blair J., Czaja R, Blair E. (2014), Designing Surveys. A guide to decisions and procedures, Sage

Couper, M.P. (2000), 'Web Surveys: A Review of Issues and Approaches' *Public Opinion Quarterly*, 64 (4): 464-494

Fischer M., Wang J. (2011), Spatial data analysis. Models, methods and techniques, Springer

Florida R. (2003), The Rise of the Creative Class, New York, Basic Books

Florida R. (2005), Cities and the Creative Class, New York, Basic Books

Ganzeboom H. (1988), Leefstijlen in Nederland: Een verkennende studie . Sociaal Cultureel Planbureau, Rijswijk.

Glass R. (1964), 'Introduction: Aspects of change' in *London Aspects of Change*, Centre for Urban Studies (ed.)

Groves R. M. (2004), Survey Errors and Survey Costs, Paperback

Hamnett C. (1991), 'The blind men and the elephant: the explanation of gentrification', *Transactions of the Institute of British Geographers* 16 (2): 173-189

Hamnett C. (1994), 'Socio-economic change in London: Professionalization not polarization', *Built Environment* 20 (3): 192-203

Hamnett C. (2003), 'Gentrification and the middle-class remaking of inner London 1961 -2001*, Urban Studies* 40 (12) : 2401-2426

Kalton G., Flores-Cervantes I. (2003), 'Weighting methods', *Journal of Official Statistics* 19 (2): 81–97

Lees L, Slater T., Wily E. (2008), Gentrification, Routledge

Singer E., Ye C. (2013), The Use and Effects of Incentives in Surveys, *The Annals of the American Academy of Political and Social Science* 645: 112-141

Smith N. (2002), 'New globalism, new urbanism: Gentrification as global urban strategy', *Antipode* 34 (3): 427-450

Sweeney S.H, Konty K.J. (2004), 'Robust point pattern inference from spatially censored data', *Environment and Planning*

Van Acker (2010) 'Spatial and Social variations in Travel Behaviour. Incorporating lifestyles and attitudes into travel behaviour-land use interaction research', doctoral thesis, Department of Geography, Faculty of Sciences, Ghent University

Tourangeau R., Conrad F.G., Couper M. P (2013), The science of Web Surveys, Oxford University Press

Van weesep J. (1994), 'Gentrification as a research frontier', *Progress in Human Geography* 18: 74-83

Chapter 3

Data cleaning and screening

3.1. The sample

The launch of the online questionnaire was set on October 12: it was published on ACTA[1] website (www.actainrete.it). No pre-notice email and reminders were sent as no contact information were available.

The survey was available online for a total of 10 weeks between October and December 2015.

The questionnaire was closed on December 24 and the number of total respondents was 573 (400 valid responses, 173 incomplete ones).

The survey task involved responding to questions related to working condition, socio-economic status, residential condition, ICTs use, mobility and lifestyle in 6 sections, for a total of 86 items.

Satisfying response rate was reached thanks to

- different ways of advertising the survey: taking part to the 'freelance day' in Torino (October 28), leaving leaflets in several co-working in Torino, Milano and Genova (in November)
- lottery incentives[2] : 5 prepaid Feltrinelli gift-cards were randomly assigned to 5 respondents among the 400 who completed the whole questionnaire[3].

The web tool used for data collection is Limesurvey, made available by Università del Piemonte Orientale.

Some diagnostics, such as responding trends within the 10 weeks, were also collected with Google Analytics.

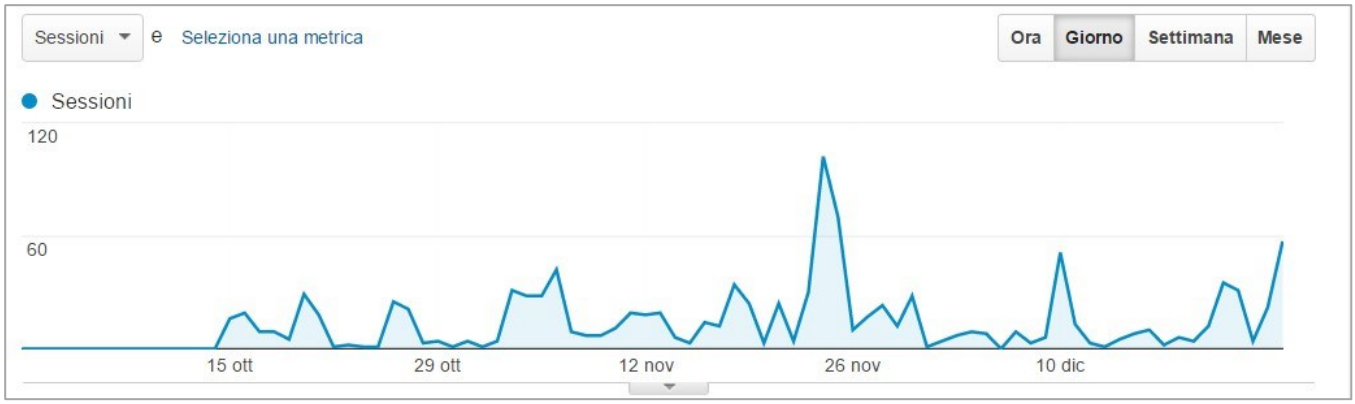
[1] An Italian quasi-union association born in 2004 in order to represent the interests of freelancers working in the advanced tertiary sector, part of the European Forum of Independent Professionals (EFIP)

[2] "Both theory and observation confirm the importance of incentives (including but not limited to monetary incentives) for participation in surveys" (Singer, Ye 2013).

[3] There was no obligation for participants to enter the lottery: they had to fill in a separate field with email contact.

Graph 1 shows the number of sessions per day.

Graph 1: Number of sessions per day



(*source: Google Analytics*)

Graph 2 shows for each session whether it was opened by a new visitor or a returning visitor. Indeed the respondents were given the possibility to leave the session (partial results were automatically saved) and to enter again and complete the questionnaire later.

Graph 2: Percentages of new and returning visitors

(source: Google Analytics)

Detected bounce rate[4] was 60,14%

[4] Bounce rate is the measure of the effectiveness of a website to retain visitors: "it is expressed as a percentage and represents proportions of visits that end on the first page of the website" (Wikipedia).

3.1.1 *Weighting and post-stratification adjustment*

Non-response is a common bias for web surveys and may cause some groups to be over- or under represented[5]: differences in the inclusion probabilities co-vary with the survey variable of interest.

Post-stratification weights are actually a useful tool to reduce selection bias and generalize findings from a sample to a larger population. A commonly applied correction technique is weighting adjustment: it assigns an adjustment weight to each survey respondent so that the total weight for each sample stratum is the same as the stratum size in the population.

As we deal with survey data concerning the working activity we made reference to Italian Labour Force Survey (2011) in the attempt to quantify the population of interest.

We first filtered the whole sample of the national survey by attributes[6] reflecting main characteristics of our target population (Ipros), then we applied expansion weights in order to compute frequencies so that the weighted data reflects the size of the population.

Post-stratified variables[7] used to compute frequencies are gender, education and age.

As shown from *Table 1*

- cell percentages are low when considering the distinction between “highly specialized professionals” and “higher level technicians” (ISCO08)
- cell percentages are very low (<1) for female over 55

So that the distinction based on ISCO classification was not considered and age classes ’55-64’, ’65-74’ ‘75+’ were recoded in one single class ‘over 55’.

[5] By comparing the observed frequency distribution of a variable with its population distribution, you can establish whether the survey response is representative with respect to this variable. If there substantial difference between the response distribution and the population distribution, you can draw the conclusion that there is a lack of representativeness with respect to this variable.

[6] As listed at page 36 and 37:
- independent working time (C6=1)
- independent working place (C7=1)
- being a collaborator or self-employed (DIPAUT=2 | DIPAUT =3)
- being a professional or own-account worker (POSPRO=8|POSPRO=9)

[7] At first also the geographic area was included (North, Centre, South), but marginal percentages became too small by considering this stratum, too.

Table 1: High specialized professionals and higher level technicians by gender, age and educational level

				Highly specialized professionals		*Higher level technicians*		*total*	
			Age	*a.n*	*% by cell*	*a.n*	*% by cell*	*a.n*	*% by cell*
	Male	*Tertiary education*	*25-34*	37.654	4,3	10.157	1,2	47.811	5,5
			35-44	90.920	10,5	17.124	2	108.044	12,5
			45-54	59.000	6,8	6.630	0,8	65.630	7,6
			55-64	47.801	5,5	5.621	0,6	53.422	6,1
			65-74	17.236	2	1.422	0,1	18.658	2,1
			75 +	5.106	0,5	629	~ 0	5.735	0,5
		tot		257.717	**30**	41.583	4,8	299.300	**34,6**
		Secondary education	*25-34*	5.313	0,6	48.409	5,6	53.722	6,2
			35-44	7.780	0,9	87.113	10	94.893	10,9
			45-54	12.775	1,5	78.072	9	90.847	10,5
			55-64	6.846	0,8	41.712	4,8	48.558	5,6
			65-74	1.448	0,1	9.878	1,1	11.326	1,2
			75 +	626	~ 0	1.032	0,1	1.658	0,1
		tot		34.788	**4**	266.216	**30,8**	301.004	**34,8**
Tot	600.304	600.304		292.505	**33,8**	307.799	**35,6**	600.304	**69,5**
	Female	*Tertiary education*	*25-34*	34.353	4	8.363	0,9	42.716	4,9
			35-44	78.006	0,9	7.210	0,8	85.216	1,7
			45- 54	31.697	3,6	4.404	0,5	36.101	4,1
			55-64	9.240	1	747	~ 0	9.987	1
			65-74	3.739	0,4	225	~ 0	3.964	0,4
			75 +	346	~ 0	-		346	0
		tot		157.381	**18,2**	20.949	**2,4**	178.330	**20,6**
		Secondary education	*25-34*	3.607	0,4	14.613	1,7	18.220	2,1
			35-44	5.447	0,6	20.550	2,4	25.997	3
			45- 54	6.567	0,7	20.676	2,4	27.243	3,1
			55-64	3.453	0,4	7.809	0,9	11.262	1,3
			65-74	327	~ 0	1.867	0,2	2.194	0,2

			75 +	262	~ 0	-		262	0
		tot		19.663	**2,2**	65.515	**7,6**	85.178	**9,8**
Tot	263.508	263.508		177.044	**20,5**	86.464	**10**	263.508	**30,5**
Total	863.812	863.812		469.549	**54,3**	394.263	**45,7**	**863.812**	100

(Source: own elaboration on Istat, LFS 2011)

Post-stratification was then used to adjust the sampling and replicate frequencies so that the joint distribution of a set of post-stratifying variables fairly matches the known population joint distribution.
In order to proceed this way and assign a weight to each record, respondents with missing values for any of the stratifying variables were deleted from the dataset: then final analysis involved a subset from the original dataset made of 400/573 respondents (173 records were excluded).
To summarize, the final weights were developped using the following steps:

- Applying expansion weights to the Sample of LFS, so that the weighted data from LFS reflects the size of the population
- Computing frequencies on weighted LFS (auxiliary variables are gender, age and education)
- Comparing frequencies from LFS with own collected data
- Calculating post-stratification weights[8] for own collected data

Final weighted survey data resulted as follows in *Table 2*. Some respondents, despite weighting adjustments, are still overestimated (females 25-45 with tertiary education) or underestimated (males over 55 and females w35-45 with secondary education): this is due to the fact that weights were trimmed.

Table 2. Frequencies of respondents on weighted survey data by stratified variables, % by cell

	Age	*Secondary*	*Tertiary*
Male	*25-35*	5,4	5,7
	35-45	13,7	10,4
	45-55	8,2	10,8
	over 55	3,1	4,7

[8] When the weights were too small or too large (rule of thumb: bounds are no smaller than 0.3 or larger than 3) we trimmed the weights.

Female	*25-35*	5,2	0,79
	35-45	7,0	5,1
	45-55	4,6	3,1
	over 55	1,5	2,8

(own collected data, R elaboration)

3.2. Data cleaning

In any statistical analysis it is very important to have clean and reliable data so not to get to misleading results. The whole process of data cleaning involved different checks on data varying from simple recoding to data screening methods.

Preliminary checks involved illegal values, misfielded values (i.e. postal code and Town were inverted), contradicting records, spell-checking and standardization (for string variables).

Later missing values and outliers were detected and data screening methods were implemented in order to identify respondents who have failed to provide honest or thoughtful responses.

At last the process of data editing also involved the creation of new variables and indexes.

3.2.1 *Missing values and outliers*

A first check involved variables with high number of missing data.
Exception was made considering filtered variables (which do not apply to all the respondents[9]), and variables were multiple answers were needed ('select all that apply' instruction).

Mainly 3 variables have been deleted from the dataset

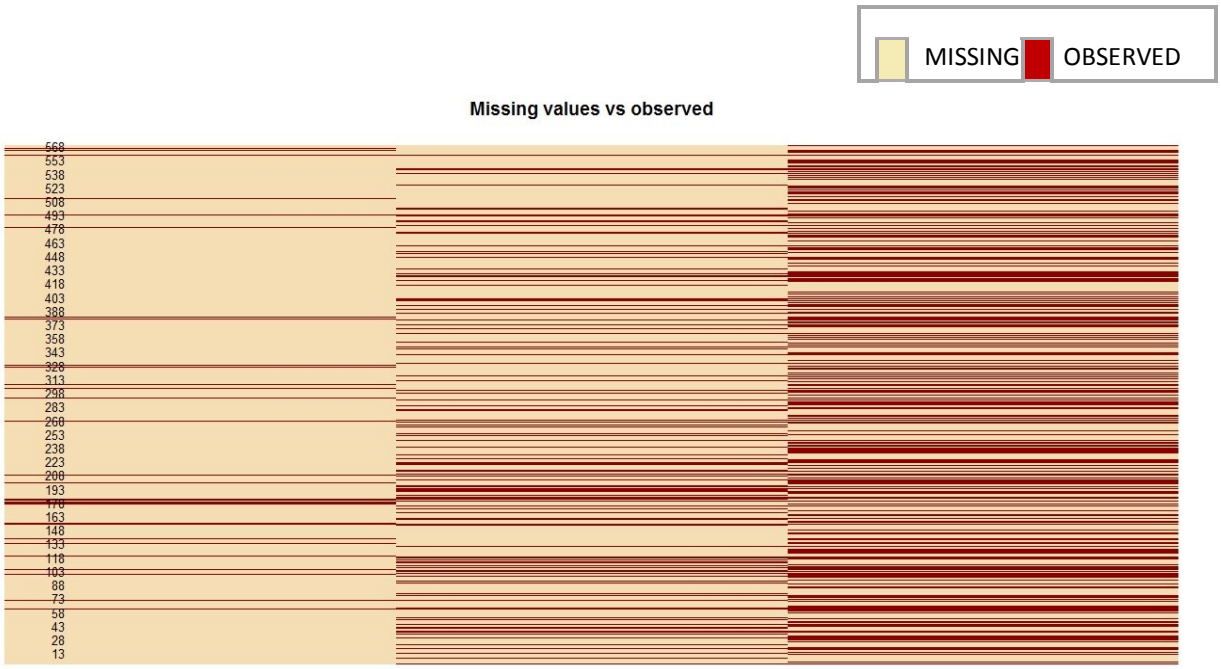
A051: "do you pay for any complementary pension program?"

C113: "do you have family strategies to face your job-related instability ?"

E094: "do you often use high-speed train for job-related mobility?"

[9] I.e: if the respondent have answered 'YES' to the question 'Do you still live with your parents' (var B04), it is nonsense to ask him/her 'When did you move in your current dwelling ?' (var B041)

Graph 3: variables with higher number of missing data



A051= 546/573 C113= 447/573 EO94= 324/573

(source: own collected data, R elaboration)

Response time for each respondent was then considered as a first 'screening technique': "Using response time as a screening technique relies on the assumption that there is a minimum amount of time that respondents must spend on an item in order to answer accurately" (DeSimone et al., 2015: 173).

Despite respondents vary in reading speed and different items have different length, too fast respondents may be problematic: checking responding time is effective for detecting careless responders; people who spent less time completing a survey than reasonably would be considered careless respondents.

The literature consider minimum responding time from 2 seconds up to 5: according to Huang et al. (2012 : 106) "it is unlikely for participants to respond to survey items faster than the rate of 2 s per item", whereas Curran et al. (2010) used a cut point of 5.5 seconds per item.

Computer-administered survey, as this is the case, offer more precision measuring the amount of time spent by each respondent on the survey: the variable "answering time", automatically calculated and provided by the software Lime Survey, shows evidence of very high responding times but no low ones.

If we consider as cut-point 5.5 seconds per item we easily find out that (5.5 * 86)= 473 seconds, which is fairly below the minimum of 603 (around 10 minutes) as from the summary in *Table 3.*

Table 3: Minimum, maximum and mean of 'responding time'

Descriptive Statistics for answering time (unit of measure: seconds)

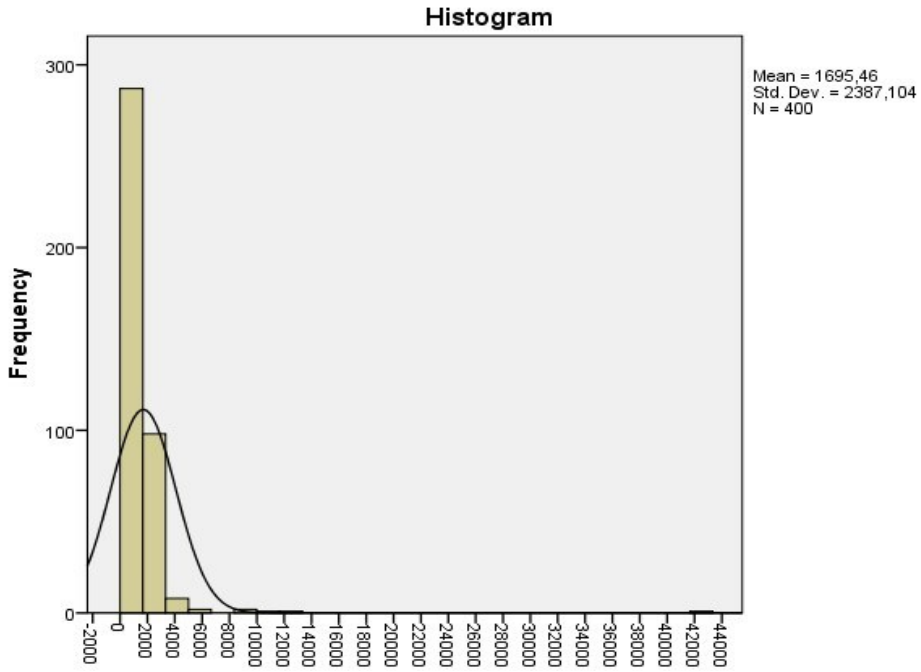
Variables	N	Minimum	Maximum	Mean	Std. Deviation
Responding time	400	603	43096	1695,46	2387,104
Valid N (listwise)	400				

(source: own collected data, SPSS elaboration)

On the other hand there are respondents who took an extremely long time to finish the questionnaire[10]. Reasons which would suggest not to consider these cases may be related to the fact that participants, by interrupting the answering process, lost concentration and could have provided careless answers as well as those who answered too fast.

Distribution of frequencies (graph 4) suggests not to consider respondents who took more than 8500 seconds (142 minutes) to complete the questionnaire.

Graph 4: Distribution of frequencies of 'responding time'

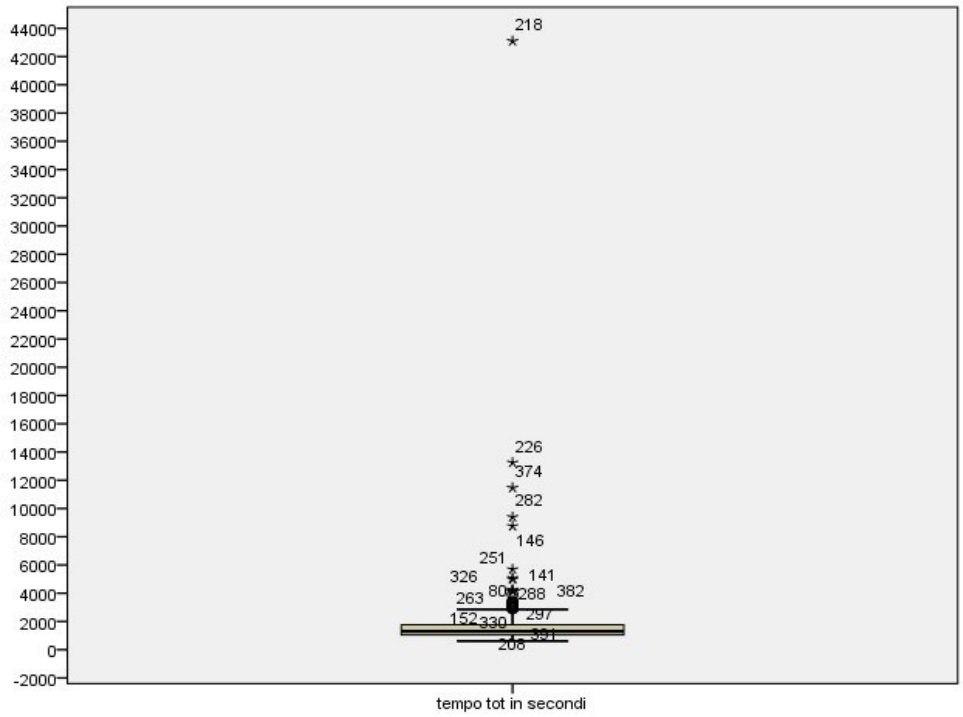


(source: own collected data, SPSS elaboration)

The boxplot (graph 5) clearly identifies the 5 outliers by case number, whose values are available in *Table 4*.

[10] This effect was also pushed by the choice to allow people to leave the session and enter again and complete the questionnaire later.

Graph 5: boxplot of 'answering time' with case number



(source: own collected data, SPSS elaboration)

Table 4: 'Answering time' top-5 highest values

Extreme Values

		Rank	Case Number	Value
Answering time	Highest	1	218	43096
		2	226	13245
		3	374	11464
		4	282	9395
		5	146	8751

(source: own collected data, SPSS elaboration)

Final number of valid records was then set to 395 (400-5 respondents).

3.2.2 Data screening methods

In any type of survey research, inattentive or careless responses are a concern: such data could lead to spurious within-group variability and lower reliability, which in turn will tend to attenuate correlations (Clark, Gironda, & Young, 2003: 1).

Response set[11] and opposite responses are both a clear symptom of response bias, where the source of the error is due to inattentive and careless respondent's behaviour.
Fortunately response bias could be detected by applying statistical screening methods which "require no survey modification and rely on statistical techniques" (DeSimone et al., 2015: 172). By using two different data screening methods we intend to identify respondents who have failed to provide honest or thoughtful responses.

First we have identified most similar survey items chosen among those used to built indexes (see par. 2.3) with same 5 or 8 points Likert scale. Among these items the phenomenon of 'response set' can easily be checked for by computing standard deviation (sd) by row: subjects with lowest standard deviation (sd $\cong$ 0) are those having chosen always the same answer option.

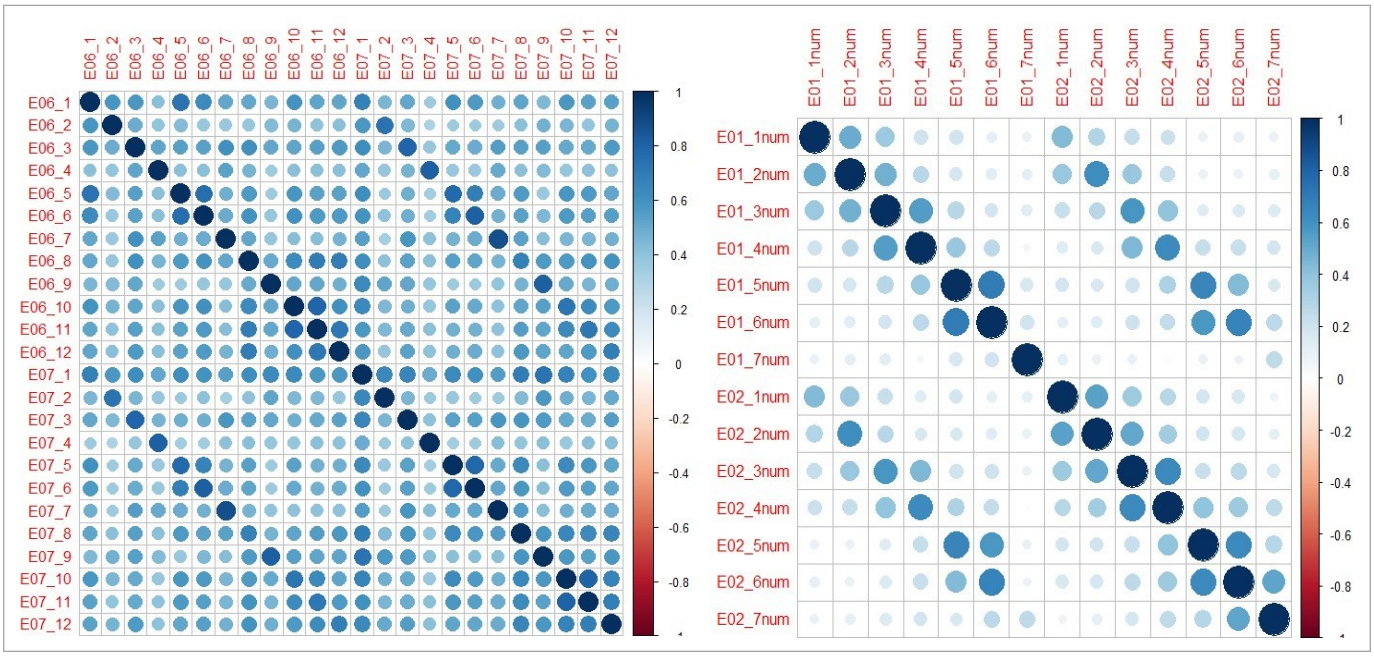
The first index (14 variables from E01_1 to E02_7) showed no subjects with sd=0
The second index (24 variables from E06_1 toE07_12) showed no subjects with sd=0

In order to detect the opposite phenomenon we relied on the construction of consistency indexes, formed by examining the differences in two items that are highly similar or dissimilar in content. The psychological literature (Meade and Craig 2012 , Goldberg 2000, Johnson 2005) talks about Psychometric Synonims and Psychometric Antonyms respectively.

These synonym /antonym screening indices are all computed in a similar manner: the first step is to identify item pairs by examining the inter-item correlation matrix.
Correlations among most similar survey items among those listed to build the 2 indexes have been computed and 16 item pairs with the largest significant positive correlation[12] ($r > 0.70$[13]) were identified. The correlation graph allows to immediately visualize the most correlated variables (darker blue circles).

[11] Response set phenomenon is a systematic tendency for respondents or subjects in tests, questionnaires, etc., to respond in a positive manner or to give the same answer to multiple-choice questions, despite their true beliefs (Oxford Psychology Dictionary). They tend to minimize their effort.
[12] Psychometric Antonyms is formed in the same way as the Psychometric Antonyms except that item pairs with the largest positive correlations are used as the relevant item pairs.
[13] The criterium also chosen by Meade et al. (2011) is $r>0.60$ ensures items are sufficiently similar in meaning.

Graph 6: Plot of correlation matrix among 2 sets of variables



(source: own collected data, R elaboration)

Once identified the item pairs, "the researcher must correlate the vector of responses to the first items in the set with the vector of responses to the second items in the set" (DeSimone 2015: 175): this correlation can be computed for each respondent and serves as the screening index (psychometric synonym coefficient).

We first filtered the 395 respondents by selecting the most problematic with highest sd (index 1, 8- point Likert scale, sd = 3; index 2, 5 point Likert scale, sd = 1.8): only for these respondents the synonym coefficient was measured. Effortful respondents should have high (positive) values for the coefficient: Meade and Craig (2012) eliminated respondents with psychometric synonym coefficients below 0.22, so we also set the cut point at similar level, 0.3.

On the overall the following records were deleted:

- No subjects with sd = 0 and correlation = 1
- 1 subjects with synonym coefficient below 0.

The total amount of final valid records was set at 394.

3.2.3 Indexes building

An index is an accumulation of scores from a variety of individual items and is often a useful measure of multidimensional and latent phenomena.

Content validity (do the index really represent the phenomenon we want to study?) has been taken into account with a specific reference to previous surveys in the field of Mobility studies.
The choice of the dimensions and of the indicators has been led by previous empirical studies such as ISTAT Census (2011), JOB MOB (2007) and the Study of Mobility in the Netherlands (OViN 2013)[14].

Reliability of indexes and their scales has been checked by calculating Cronbach's α coefficient which is higher where items/indicators are higher correlated one to each-other[15].

Here a quick overview of the features of each index (including type of scaling, range of scores, possible weights of each item[16] and Cronbach's α):

- **Index of mobility** relies on the aggregation of indicators in multidimensional terms: first dimension is represented by the frequency of travelling, whereas the second by the distance travelled.
  A score is assigned to people travelling more than once a month in areas in ascending order (from the smallest 'the neighbourhood', to the greatest 'other Italian regions'[17]):

Table 5: Mobility index: indicators, mean and standard deviation, Cronbach's Alpha reliability coefficient

ITEM	SCALE/ CRITIERIUM	WEIGHT/ SCORE
E01: job-related mobility in neighbourhood	>= 'more than once a month'	1
E01_2: job-related mobility in town	>= 'more than once a month'	2
E01_3: job-related mobility in metrop. area	>= 'more than once a month'	3
E01_4: job-related mobility in the region	>= 'more than once a month'	4
E01_5: job-related mobility in adj. reagions	>= 'more than once a month'	5
E01_6: job-related mobility in other regions	>= 'more than once a month'	6

[14] Among others some measures were collected such as: distance between home and workplace, displacement time, most used means of transportation, urbanity class, number of trips, number of occupational displacements per day.
[15] A basic criterion for choosing which items to include in an index is uni-dimensionality. That is, each item should only represent only one dimension of the concept you are measuring. However items used to build a specific index should also be empirically related: "if two items are empirically related to each other, we can argue that both items reflect the same concept".
[16] It is possible to weight individual items according to their importance in the concept being measured.

Indicator	Mean	St.Dv	Cronbach's α
Mobility_index	6.5	1.3	0.7

(source: own collected data, R elaboration)

- **ICTs use synthetic index** was built by adopting a different cumulative strategy of aggregation of the indicators

Table 6: ICTs use index: indicators, mean and standard deviation, Cronbach's Alpha reliability coefficient

ITEM	SCALE/ CRITERIUM		WEIGHT/ SCORE
D01: number of used IT devices	Unit Count (min. 1- max. 6)		1 - 0.5
D02: number of used networks	Unit Count (min. 1- max. 6)		1 - 0.5
D03_2: on the telephone	Each day (3) – Never (0)		1 - 0.5
D03_3: on Skype	Each day (3) – Never (0)		1 - 0.5
D03_4: by email	Each day (3) – Never (0)		1 - 0.5
D03_5: by sms / whatsapp	Each day (3) – Never (0)		1 - 0.5
D03_6: by social network	Each day (3) – Never (0)		1 - 0.5
Indicator	**Mean**	**St.Dv**	**Cronbach's α**
Tech.use_index	9.7	6.6	0.59

(source: own collected data, R elaboration)

- ISEI (International Socio-Economic Index of occupational status) for mother and father respectively was built relying upon ISCO08 indicator [17] (2 digits). Harmonization and measurement tools were provided by Ganzeboom's ISMF Project [18].

[17] The index attributes a scale of values ranging from 16 to 90. The ISCO variable was computed by analysing variables (C06_1:C06_4) concerning sector of economic activity and occupational skill level for both mother and father of the respondents. Doubts concerning self-assessment definition were solved by looking at the educational level of the parent (C05).

[18] SPSS syntax modules ("Assignment of ISEI scores to ISCO-08 codes") are available at http://www.harryganzeboom.nl/isco08/index.htm

- **Living place-related satisfaction** was built taking into account opinion of respondents concerning several dimensions ranging from social and family life to cultural, economic environment.

Table 7: Living Place-related satisfaction: indicators, mean and standard deviation, Cronbach's Alpha reliability coefficient

ITEM	**SCALE/ CRITERIUM**		**WEIGHT/ SCORE**
B043: Satisfaction infrastructure network	1 (not at all) -10 (completely)		N= non NA
B045: Satisfaction greenery	1 (not at all) -10 (completely)		N= non NA
B047: Satisfaction closeness to work	1 (not at all) -10 (completely)		N= non NA
B049: Satisfaction living cost	1 (not at all) -10 (completely)		N= non NA
B0492: Satisfaction business environment	1 (not at all) -10 (completely)		N= non NA
B0494: Satisfaction closeness to parents	1 (not at all) -10 (completely)		N= non NA
B0496: Satisfaction cultural environment	1 (not at all) -10 (completely)		N= non NA
B0498: Satisfaction social life	1 (not at all) -10 (completely)		N= non NA
B04991: Satisfaction glamour- respectability	1 (not at all) -10 (completely)		N= non NA
Indicator	**Mean**	**St.Dv**	**Cronbach's α**
Place.satisf_index	6.6	1.5	0.77

(source: own collected data, R elaboration)

3.3. Explorative descriptive results

The results here presented are taken from an ongoing study and are therefore limited in scope. The purpose is to provide a first socio-demographic description of respondents and a measure of indexes which serve as a first check of research questions.

The following analysis are based on weighted post-stratified data.

According *Table 8* and *Table 9* around three fifths of the respondents are males and over half of the respondents are aged under 45.

Crossing these data, the most represented segment of the population are men aged 35-45 (*Graph 7*). However women are more numerous than men between 25 and 35. This age range is usually the one where women fertility rates are higher and this could be a possible reason for women to choose to work as freelancers: this trend is usually also confirmed by the literature which recognises that family care still plays a big role in pushing women towards self-employment and its more flexible working conditions (Beaucage et al.2004)

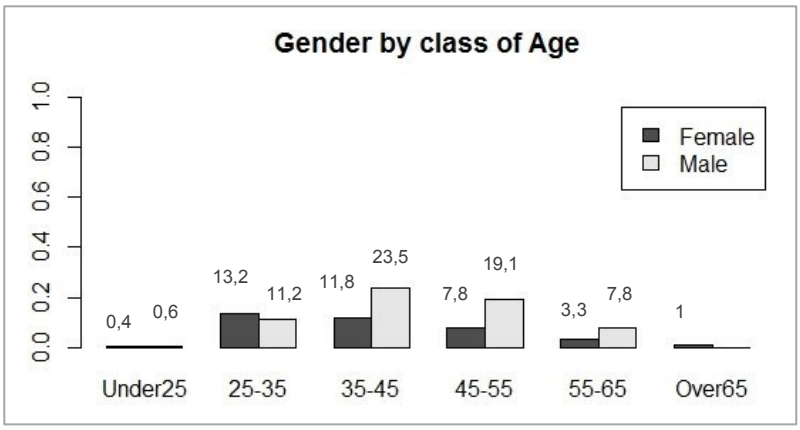
Table 8. Gender distribution (percentages)

Females	Males	*Tot*
38	*62*	100

Table 9. Age distribution (percentages)

< 25	25 - 35	35 - 45	45 - 55	55-65	> 65	*Tot*
0,9	*24,4*	*35,4*	*26,9*	*11,2*	*1*	100

Graph 7. Barplot: Class of Age by Gender



Concerning the economic dimension some data are displayed in *table 10*, each cell containing weighted percentages of the data.

Table 10. Annual gross income by education and gender (cell %)

		Annual gross income in € (cell %)					
Gender	*Education*	*<18k*	*<25k*	*<35k*	*<50k*	*>50k*	*TOT (%sum by row)*
F	Compulsory		≅ 0				0
	Diploma	7.43	4.27	2.15	3.52	1.15	18.52
	Bachelor	2.12	1.42	0.55	1	0.5	5.59
	Master	2.12	2.67	2.44	2.2	2	11.43
	Postgraduate	0.5	0.47	0.47	0.31	0.23	1.98
	TOT (% sum by col)	12.17	8.83	5.61	7.03	3.88	37.6
M	Compulsory	1.63	-	-	-	-	1.63
	Diploma	4.47	5.11	5.26	6.22	9.51	30.57
	Bachelor	1.43	1.11	1.62	1.1	1.74	7
	Master	1.74	1.26	2.25	4.4	8	17.65
	Postgraduate	0.8	0.51	1.19	1.3	1.62	5.42
	TOT (% sum by col)	10.07	7.99	10.32	13.02	20.87	62,3

(Source: own elaboration by using Rstudio software)

At first sight education level does not seems to affect income level especially for males: highest percentages of males (almost 50% among males) earning above 50.000 euros are Ipros who did not graduated.

The inexistance of the association between levels of education[19] and gross annual income is confirmed by the Chi-square test, which returns a very high p-value, not allowing to reject the null hypothesis (H0: "there is no association").

```
Pearson's Chi-squared test data:
gross_income and level_of_educ

X-squared = 35.6816, df = 16, p-value = 0.07277
```

Even when grouping the 5 educational levels into 2 ones (because of very low frequencies for 'compulsory' and 'postgraduate'), as 'secondary' and 'tertiary' education, the association is absent.

[19] In order to avoid empty cells in the calculations, compulsory education and income >100,000 has been omitted.

These first results, by showing low returns to education, seem to support the inflation of educational credentials hypothesis (Ballarino, Schizzerotto 2013)[20] especially for those with an upper secondary education level. Similarly to the "parking lot" argument (scarsity of employment opportunities would induce people to prolong their studies), we could interpret the choice of entering the freelancing world as a strategy to face youth unemployment.

Whereas concerning gender gaps, controlling for education level, males show on average higher income than women (i.e. among women less than 30% respondents earn above 35.000 euros per year, whereas among males this percentage rises to over 54%). The Chi-square supports the hypothesis that gender does affect income level among Ipros, whereas education doesn't[21].

```
Pearson's Chi-squared test data:
gross_income and level_of_educ

X-squared = 38.3058, df = 4, p-value = 9.123e-05
```

Logistic regression[22] has been also performed on same data, measuring the relative probability of having a gross annual income above 25.000 euros (dependent variable).
Less numerous categories such as people with compulsory education, aged under 25 or over 65 has not been included in the analysis.

Reference categories are

- being 'male'
- aged '35-45'
- holding a 'master' degree

Two models have been analized: Model 1 with same variables from table 10 and Model 2 adding age level and ISEI index computed both for the mother (isei_m) and the father (isei_f) of each subject.

[20] This is related to the imbalance between supply (educational attainments) and demand (occupational structure); in other words the number of highly educated people is higher than available upper level jobs.
[21] Distribution of educational levels among the 2 genders is very similar, so that for females there is not much difference with males in terms of educational attainment.
[22] Logistic regression computed on weighted data is not fitted by maximum likelihood criterium, so that standard measures of goodness of fit (PseudoR2, AIC) are not available.

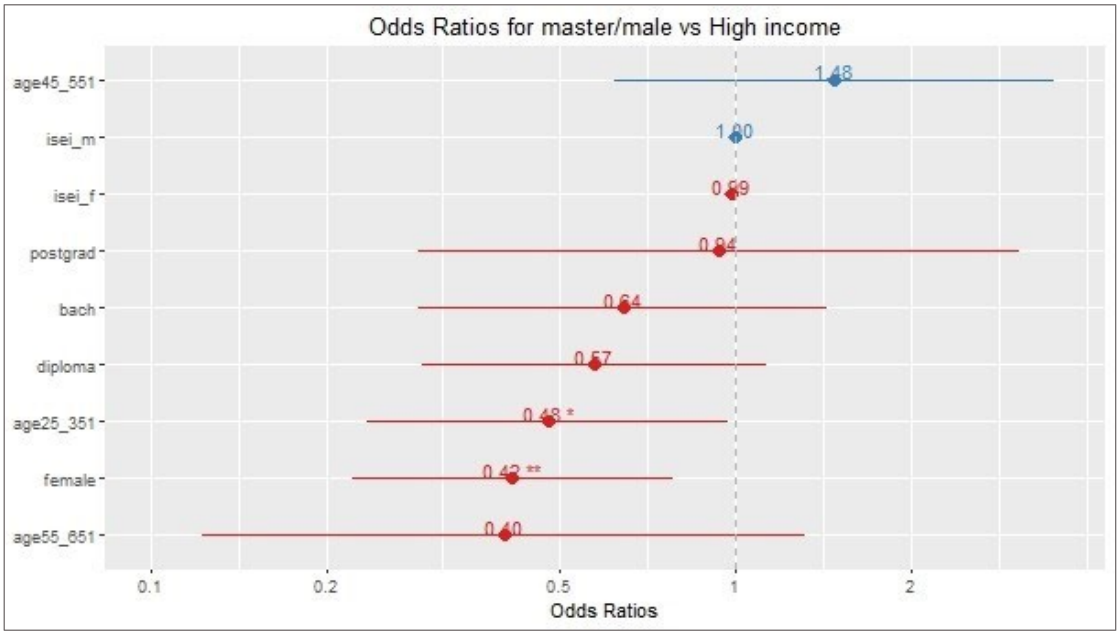
Table 11. Logistic regression models, coeff in odds.ratio
Logistic regression on gross annual income (>25.000)

	Mod I	Mod II
EDUC level (ref: master)		
Diploma	0.614	0.575
Bachelor	0.52	0.643
Postgraduated	0.949	0.937
GENDER (ref: male)		
Female	0.331 ***	0.416 **
AGE level (ref: 35-45)		
Age 25_35		0.478 *
Age45_55		1.48
Age55_65		0.402
Isei (ref: Isei_M)		
Isei_F		0.989

***p<0.01, ** p<0.05, * p<0.1
(Source: own elaboration by using Rstudio software)

To provide much more easily-readable data Odds ratios were plotted with their confidence level in *graph 8*.

Graph 8. Plot of Odds ratio estimated from Model 2 in table10



(Source: own elaboration by using Rstudio software)

Both implemented models confirm among other socio-biographic variables the highly statistically significant role of gender (α= 0.001) and the less strong role of being aged between 25-35 (α= 0.01):

According to model 2 the probability for women of earning over 25.000 euros per year is around 0.42 with respect to males with same age, educational level and family background ; the probability for youngs aged 25-35 of earning over 25.000 euros per year is around 0.48 with respect to people aged 34-45 with same gender, educational level and family background.

Surprisingly Parents' socio-economic background (ISEI index) seems not to affect this relationship. The absence of relation clearly emerges by displaying data in double-entry tables (*table 12* and *table 13*) and computing Chi-squared test.

Table 12. Annual gross income by Isei_f

	Annual gross income in € (cell %)					
Isei_f	*<18k*	*<25k*	*<35k*	*<50k*	*>50k*	*TOT (%sum by row)*
Low	3.45	1.06	3.44	2.12	2.11	12.18
Low-Medium	4.23	2.38	2.64	2.65	2.38	14.28
Medium- High	9.26	5.82	5.29	6.09	6.87	33.33
High	9.27	9.52	7.14	6.87	7.40	40.20
TOT (% sum by col)	26.21	18.78	18.51	18.00	18.76	100

(Source: own elaboration by using Rstudio software)

```
Pearson's Chi-squared test data:
gross_income and Isei_f

X-squared = 9.0759, df = 12, p-value = 0.6964
```

Table 13. Annual gross income by Isei_m

	Annual gross income in € (cell %)					
Isei_m	*<18k*	*<25k*	*<35k*	*<50k*	*>50k*	*TOT (%sum by row)*
Low	4.84	3.70	4.27	4.00	4.55	21.36
Low-Medium	5.12	4.27	3.70	3.70	4.84	21.63
Medium- High	7.12	6.55	3.42	5.12	4.56	26.77
High	7.69	5.70	6.50	5.41	4.84	30.14
TOT (% sum by col)	24.77	20.22	17.89	18.23	18.79	100

(Source: own elaboration by using Rstudio software)

Pearson's Chi-squared test data:
gross_income and Isei_m

X-squared = 5.481, df = 12, p-value = 0.94

These outcomes show that direct impact of ascripted social origines is not retrieved among Ipros and no pure occupational inheritance is detected. That is intergenerational association of inequalities might be decreasing, mediated by education and the rise of new professions.

With reference to the working status, the survey collect some data about working status held by the respondent before becoming self-employed (*Table 14*). Tracking the most recent working history of professionals allow to provide a merely descriptive test of the so-called "unemployment-push" towards self employment (Beaucage et al. 2004).

1= 'student' 5= 'unemployed'

2= 'dependent worker' 6= 'researcher'

3= ' firm owner with employees' 7= 'firm owner without employees'

4= 'atypical worker/trainee' 8= 'irregular worker'

Table 14. Frequencies of former working status

1	2	3	4	5	6	7	8	*TOT*
18	*51,4*	*1,5*	*1,7*	*23,6*	*3,4*	*0,1*	*0*	100

Over 50% of the respondents have a working history as former employees.
Nevertheless, considering students and unemployed people (not studying and not working) as one bigger category of 'not employed' workers, allow to frame the issue differently: there are more than 41% of respondents whose previous situation was *de facto* an unemployment status. These figures give their contributes in supporting the thesis of "unemployment-push" which sets unemployment as main pushing factor towards self-employment.

The data also provide a reasonable explanation concerning non-weighted age and gender distribution among respondents: Women and youngs. Indeed self-employment may in some cases represent a viable alternative to unemployment, especially for disadvantaged groups of jobseekers such as women and youngs.
With regard to the spatial dimension most of the Ipros who answered the survey are located in the polycentric North West region (including Val d'Aosta, Piemonte, Liguria,Lombardia): 46 % (that is 204 out of 440) of respondents live in this area, despite not homogenously distributed.
This data are not enough to support the idea of the existence of a Polycentric city region model, however.

Indeed very high responding rates were also detected within people living in single cities (or metropolis) such as Milan, Rome, Torino: 29,5% of respondents.

What is then the best territorial unit in order to study the phenomenon and detect cluster of spatial aggregation? We have tried to answer this question in chapter 4.

Concerning mobility dimension, two main sub-dimensions have been analysed:

1. residence mobility
2. short-term commuting

Among total number of 394 respondents 88,8% of them live no more with their family of origin and experienced residence mobility out of the parents' dwelling (considered variables are B01 VS B02).

Furthermore for this subgroup of respondents we have calculated difference (in years) between date of starting working as freelancer and date of last moving into their current dwelling: on average the moving toward another dwelling took place 1,5 years after becoming autonomous self-employed[23].

Despite this data does not allow to make any assumption concerning the association between the two phenomena, we managed to compare on descriptive basis the place of living before becoming an Ipros with the current place of living of each respondent:

- *46,2 % changed city of residence* among them
- *42,3% changed city but kept living in the same region*
- *57,7% changed region*
- *3,3% moved from Italy to abroad - 7,7% moved from abroad to Italy*
- *6% moved from Southern to Northern Italian regions*
- *3% moved from Northern to Southern Italian regions*

These percentages show high rates of long-distance moving (i.e. in other Italian regions) and unexpected consistent numbers of people moving from abroad towards Italy. Age average for these people is below 35, so it could be reasonable to think about moving abroad for studying or specializing reasons with the intention to come back and start one's own activity in Italy.

On the contrary average age of those who left Italy to move abroad is over 45. This opposite trend, though based on very small percentages, shows this category of workers to be very mobile on long distances despite the aging factor.

[23] We cannot claim the straight association of the two phenomena since it was not possible to control for other intervening variables affecting the decision of moving.

Considering the analysis of the place of living a new variable was created in order to classify the 'urbanity class'. The classification of the municipality the respondent lives in is based on the administrative function, the embeddedness in metropolitan areas and the number of inhabitants:

city = regional main centre (administrative definition)

metropolitan area = a municipality which is located within the metropolitan area of a biggest city (usually a regional main centre)

province (administrative definition)

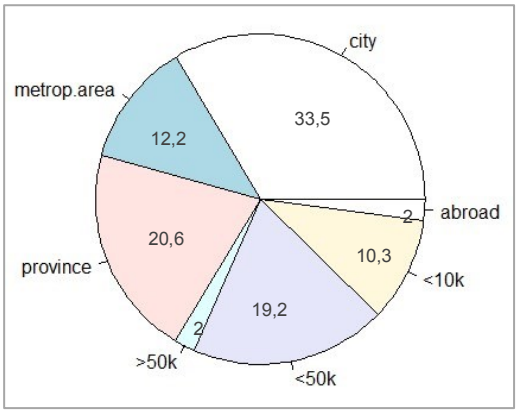
>50 k = municipality with number of inhabitants over 50.000

<50 k = municipality with number of inhabitants greater than 10.000 and less than 50.000

<10k = municipality with number of inhabitants under 10.000

Graph 9 shows frequency distribution for each category.

Graph 9. Pie plot of urbanity class



Over 45% of respondents live in a city or within its metropolitan area, and the cumulative percentage considering provinces raises to 66%. This figures provide a first proof of our hypothesis according to which Ipros prefer living in dense urban areas (urban regions) where specific economic site conditions[24] tend to promote their geographical concentration (or clustering).

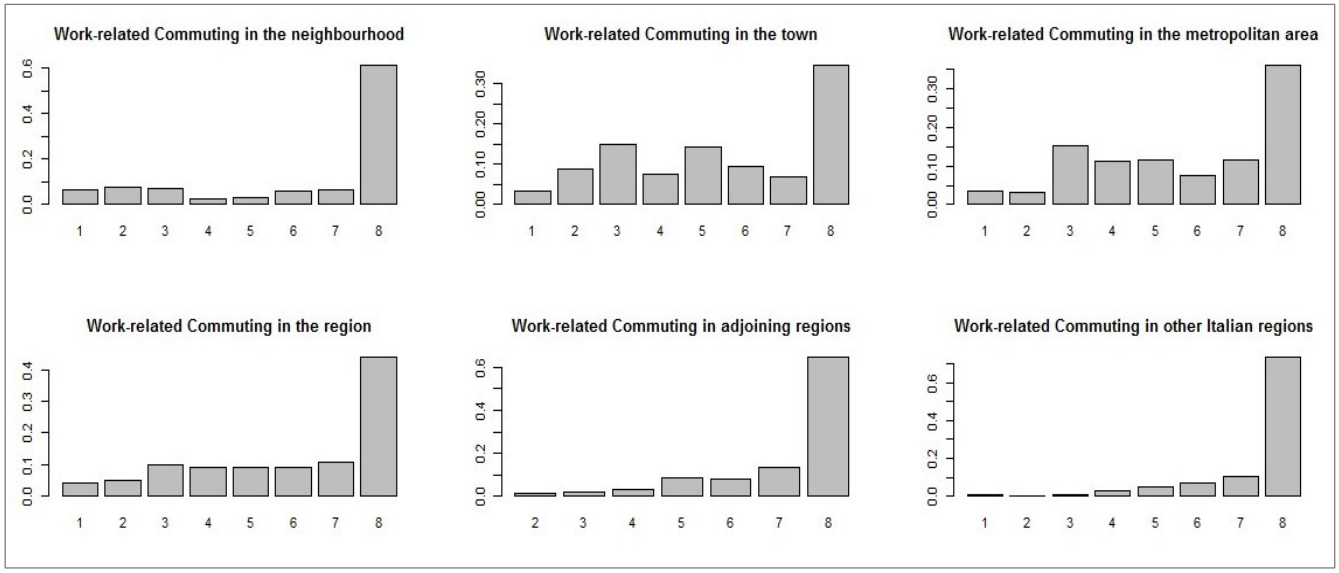
Seeking to show a general picture of people behaviour, indexes previously analysed allowed to provide measure of work-related commuting frequency (mobility index) with reference to destination and origin (urbanity class of place of living).

[24] In economic geography "central place" theory was based on similar assumptions: Goette (1994) considered economic site conditions such as sales potential, infrastructure and transportation and housing market as well as cultural ones affecting differences in locational choices.

The plots in *Graph 10* shows the number of people travelling with certain time frequency in a specific area. Areas are 6 and they are defined in terms of distance from the place of living (neighbourhood, town, metropolitan area, same region, adjoining regions, other regions), whereas time frequencies range from 1 to 8 with the following meanings:

1= 'more than once a day'
2= 'once a day'
3= 'more than once a week'
4= 'once a week
5= 'more than once a month'
6= 'once a month'
7= 'less than once a month'
8= 'never'

Graph 10. Distribution of work-related commuting frequencies by destination place

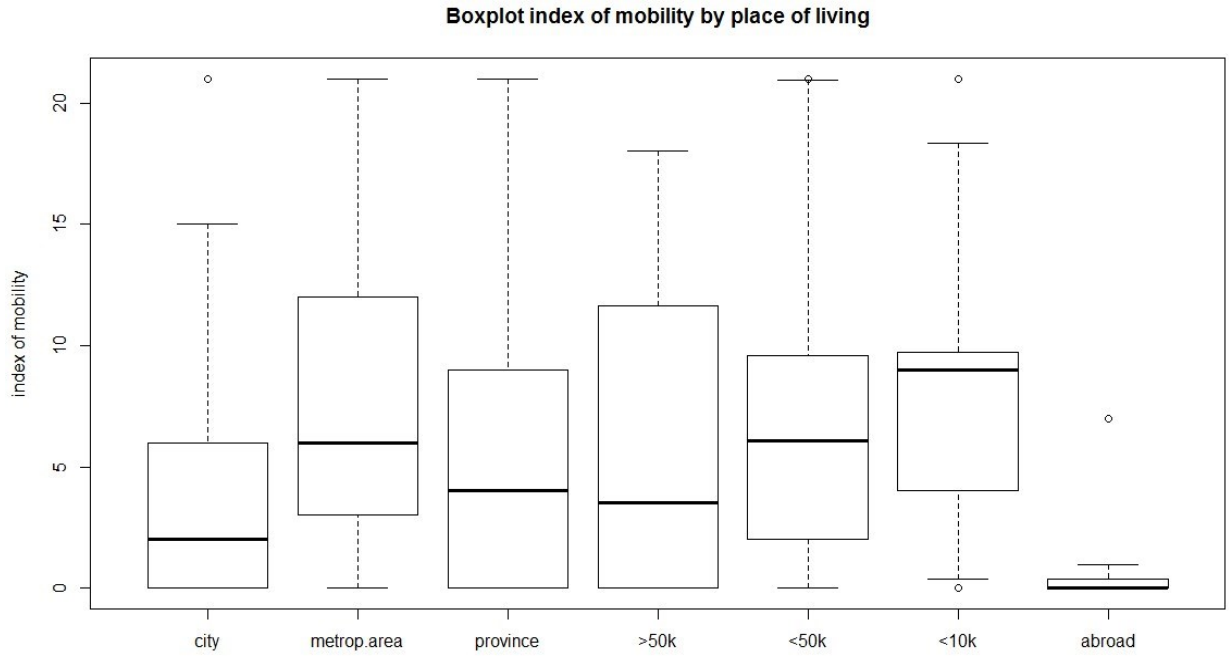


(Source: own elaboration by using Rstudio software)

Plots of data in *Graph 10* show most of the I-pros undertaking the survey are likely to move between 'more than once a week' and 'more than once a month' in areas such as the town they live in, the metropolitan surrounding area and the region.

In order to provide a clearer picture and relate the place of living to the areal distance travelled, commuting frequency variable has been matched with urbanity class categories: the following boxplot (*Graph 11*) provide a visual synthesis of the distribution of frequencies of the Index of mobility for each type of municipality the respondents live in.

Graph 11. Boxplot of Mobility index by urbanity class of place of living.



(Source: own elaboration by using Rstudio software)

As shown percentages of respondents with higher values of Index of mobility are more likely to live in smaller municipalities (>50 k , <50k , <10k). In synthesis if you are an Ipros the further you live from big centres, the more you are prone to travel long distances for work-related reasons.

The only exception is the 'metropolitan area': living in a municipality in a metropolitan area, whatever the effective dimension of the municipality, leads to quite high levels of mobility with a median of the distribution around 5. This could be explained by a rather high frequency of mobility within the metropolitan area the municipality belongs to[25]; on this purpose Martinotti (in Nuvolati 2002) talks about "third-generation metropolis" where urban sprawl has enhanced local mobility and short-range up to long-range commuting patterns.

An Anova was run in order to test this hypothesis and mean of mobility index (var. E01_index) proved to be different for each category of the urbanity class (var. B02_dim_class) with quite a high confidency level (α= 0.01)

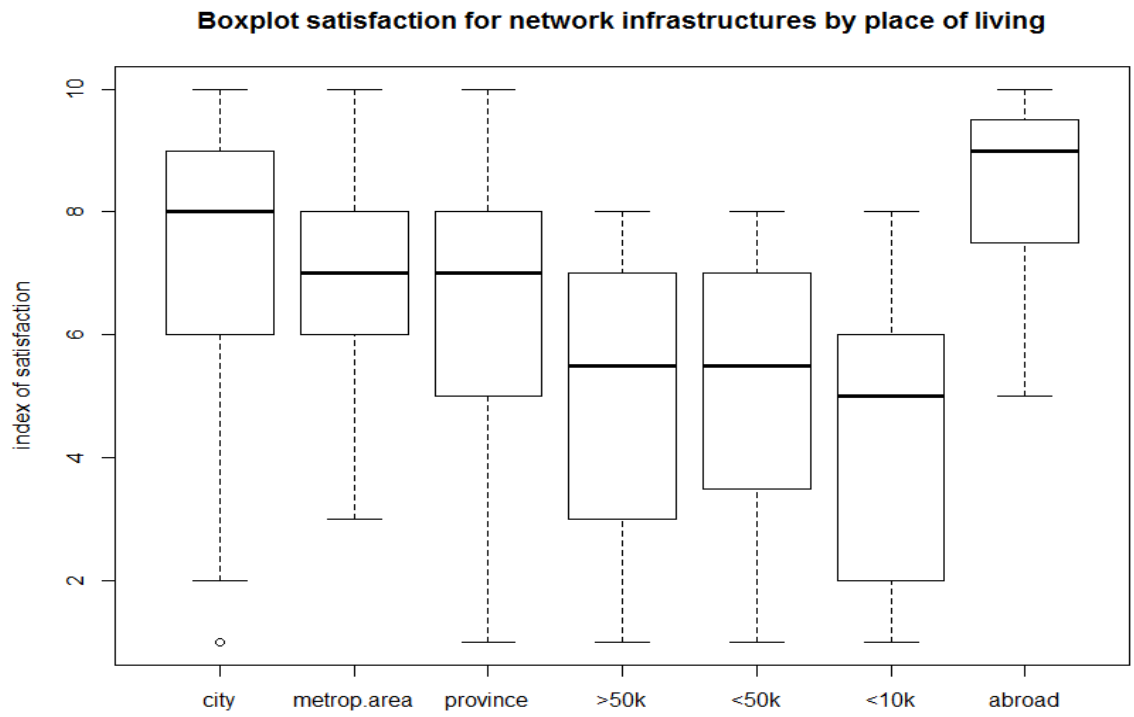
Analysis of Variance Table

```
Response: dataset$E01_index
                       Df  Sum Sq Mean Sq F value   Pr(>F)
dataset$B02_dim_class   2   348.5 174.246  6.8035 0.001232
** Residuals             436 11166.5  25.611
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

[25] urban sprawl has enhanced local mobility and short-range up to long-range commuting patterns ("third-generation metropolis").

Among indicators used to build up the living place-related satisfaction index, satisfaction for network infrastructure was a key one (var. B043_SQ001).

Graph 12. Boxplot of Mobility index by urbanity class of place of living.



(Source: own elaboration by using Rstudio software)

The boxplot (*graph 12*) and the anova computed on this indicator and the urbanity class shows a clear relation among the two: satisfaction for local infrastructures varies according to the urbanity class of place of living: the smaller the municipality the lower the satisfaction.

```
Analysis of Variance Table
Response: dataset$B043_SQ001
                       Df  Sum Sq Mean Sq F value    Pr(>F)
dataset$B02_dim_class   2  155.13  77.565  12.656 5.285e-06 ***
Residuals             300 1838.58   6.129
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Finally the probability of showing Mobility index score higher than the mean of 4.9 was computed for two nested logistic models. Reference categories are

- being ‘male’
- aged ‘25-35’
- being ’single’

Table 15. Logistic regression models, coeff in odds.ratio

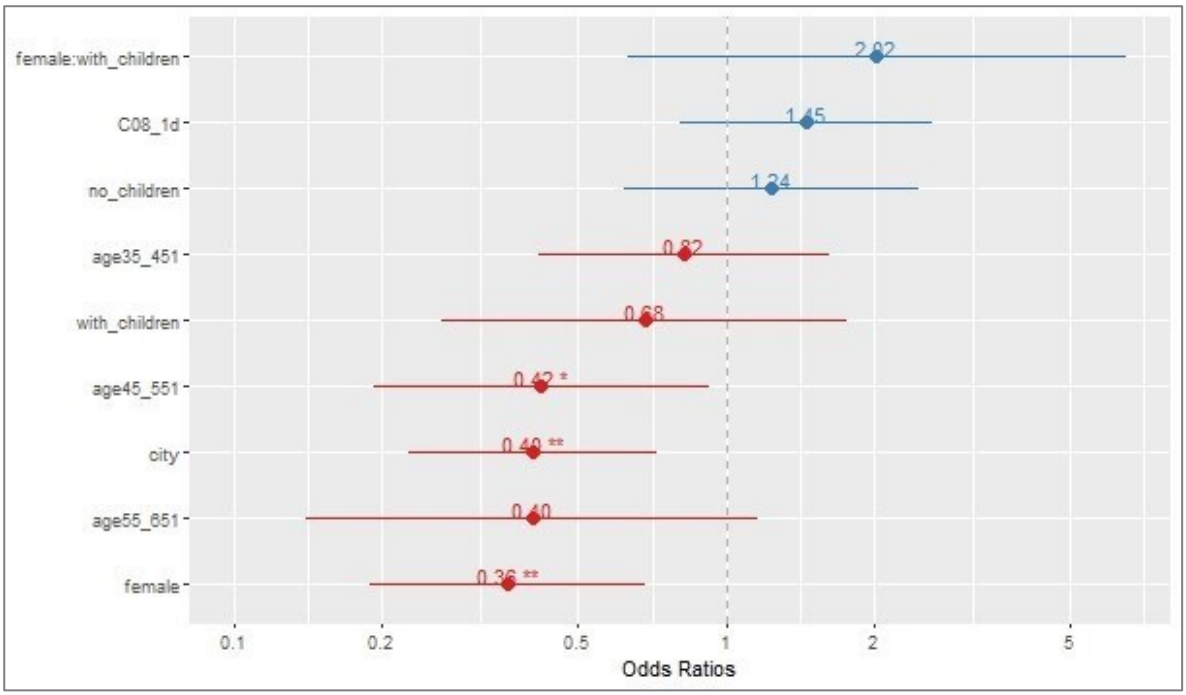
	Mod I	Mod II
FAMILY CONDITION (ref: single)		
With_children	-0.06988	0.3828
No_children	0.20723	-0.3140
URBANITY CLASS (ref: no city)		
City	-0.3140 **	-0.9057 **
GENDER (ref: male)		
Female	-0.84411 **	-1.0237 **
AGE level (ref: 25-35)	FF	FF
Age 35_45	-0.15438	-0.1994
Age45_55	-0.76739 *	-0.8643 *
Age55_65	-0.87315	-0.9087
FAMILY * GENDER (ref: Males without children)		
female with_children		0.7049
INCOME (ref: low)		
High income (C081_d)		0.3737

***p<0.01, ** p<0.05, * p<0.1

(Source: own elaboration by using Rstudio software)

Odds ratios are plotted with their confidence level in *graph 13*.

Graph 13. Plot of Odds ratio estimated from Model 2 in table15



(Source: own elaboration by using Rstudio software)

Model 1 introduces socio-demographic variables (age, gender,) plus family condition and urbanity class at city level.

As we expect living in a city proves to be a significant factor in decreasing probability of having a high mobility index. Being a women aged 45-55 also negatively affect the probability of performing a high score for mobility index.

Family dimension, instead, against any expectations seems not to be statistically significant. Model 2 inputs in the analysis an interaction factor resulting from the combination of being a woman (gender) with children (family composition) but the coefficient is still not significant.

Last variable in model 2 is the one measuring high income level: neither this coefficient seems to have an impact on the dependent variable.

The analyses described above show that the urbanity class is negatively associated with work-related commuting frequency and positively associated with satisfaction for network infrastructures; those who experience higher mobility are somehow those who display lower satisfaction about existing infrastructures. This aspect should be further investigated[26] as it could highlight the inadequacy of

[26] What is the mean of transport most often used by 'unsatisfied'people? Does the high frequency of use make the respondents more severe and strict in judging public transport?

network infrastructures outside dense urban areas, contributing to create possible phenomena of geographical marginalization for people living in smaller municipalities (<50 k).

In addition higher scores (above the mean) of work-related mobility are detected for young men (25-35) living outside cities; this same age group is the one which proved to have a lower income with respect to others (see *table 11*). We then suppose this category of people to be the one more likely to hold marginal positions in the labour market and eventually undergo spatial exclusion from economic centres.

Last hypothesis we want to test in this initial set of descriptive statistics is the relation between mobility index (var E01_index) and ICTs' use index (var. tech.use_index): do telecommunication enhance mobility (Negroponte 1991, Castells 1996)?

A simple correlation shows their positive and significant relationship (ρ=0.35)

```
Pearson's product-moment correlation
data:  datasetnoNA$tech.use_index and
datasetnoNA$E01_index t = 7.3952, df = 393, p-value =
8.604e-13
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.2597999 0.4332413
sample estimates:
      cor
0.3495113
```

So we can conclude that yes, according to our data, ICTs development "does not imply the end of the office, but the diversification of working sites for a large fraction of the population, and particularly for its most dynamic, professional segment"[27] (Castells 1996).

[27] It is noteworthy that Castells (1996) argues that, especially in the case of information technology activities, spatial proximity is a necessary material condition for the existence of 'milieu of innovation' because of the interactive nature of the innovation process. "Once established, milieux of innovation both compete and cooperate between different regions, creating a network of interaction that brings them together in a com- mon industrial structure beyond their geographical discontinuity."

References

Babbie, E. (2001), *The Practice of Social Research: 9th Edition*, Belmont, CA, Wadsworth Thomson.

Ballarino, G., Scherer, S. (2013). "More Investment – But Less Returns? Changing Returns to Education in Italy Across Three Decades", *Stato e Mercato, 99*: 359-388.

Castells M. (1996),*The information Age: Economy, Society and Culture*, Vol. I, Oxford, Blackwell

Clark R.E., Gironda R.J., Young R.W (2003),"Detection of back random responding: effectiveness of MMPI-2 and personality assessment inventory validity indices", *Psychological assessment 15(2)*: 223234.

DeSimone, Justin A.; Harms, Peter D and DeSimone, Alice J. (2015), "Best practice recommendations for data screening", Management Department Faculty Publications. Paper 124.

Goette T. (1994), *Location policies of international companies*, Wiesbaden, Gabler

Goldberg, L. R., & Kilkowski, J. M. (1985) "The prediction of semantic consistency in self-descriptions: Characteristics of persons and of terms that affect the consistency of responses to synonym and antonym pairs", *Journal of Personality and Social Psychology, 48*: 82-98.

Huang, Curran, Keeney, Poposki & DeShon (2012), "Detecting and deterring insufficient effort responding to surveys*", Journal of Business and Psychology 27*: 99-114.

ILO (2012), *ISCO–08 Volume I*, *International Standard Classification of Occupations Structure, group definitions and correspondence tables*

Johnson, J. A. (2005), "Ascertaining the validity of individual protocols from web-based personality inventories", *Journal of Research in Personality, 39(1)*: 103-129.

Lumley T . (2004), *Analysis of complex survey samples*, Department of Biostatistics University of Washington

Meade, A. W., & Craig, S. B. (2011, April), "Identifying careless responses in survey data", Paper presented at the 26th Annual Meeting of the Society for Industrial and Organizational Psychology, Chicago, IL.

Nunnally J. C. (1967), *Psychometric theory*, New York, McGraw Hill

Nuvolati G. (2002), *Popolazioni in movimento, città in trasformazione. Abitanti, pendolari, city-users, uomini d'affari e flâneurs*, Bologna, Il Mulino

Singer E., Ye C. (2013), The Use and Effects of Incentives in Surveys, *The Annals of the American Academy of Political and Social Science* 645: 112-141

Chapter 4

Macro-economic factors influencing Ipros locational choices: Spatial clustering detection and regression models.

4.1. Introduction and recall of background literature

The "central place theory" from economic geography describes settlements and urban nodes according to a hierarchy, underlying the "tendency for regional city space to revolve around a single dominant urban core" in the metropolitan region (Soja 2000: 16).

What if we reverse this geographical perspective into a labor market one? Do central urban nodes attract Ipros? Do they live in this areas?

Chris Hamnett (1991) provides a useful literature background (see chapter 1): he analysed gentrification process in the context of the new postindustrial city marked by "deeper changes in the structure of production, the changing division of labour, and the rise of a locationally concentrated service class" (Hamnett 1991: 177). Hamnett's thesis is also known as the "professionalization thesis" as he argued for "the increase proportion of professional workers in a number of large cities with a strong financial/producer service base" (Hamnett 1994: 407).

Our purpose in this chapter is to test this hypothesis mutated from the economic geography field by first defining "central places" according to a set of available structural variables at local level and later modelling these data in order to see whether we can get any evidence of Ipros clusterization in these areas.

We have first defined and collected some infrastructural macro variables defining the so-called "functional space" (Kauffman 2011: 52) and granting the suitable receptiveness of Ipros needs.

These data were then matched with answers collected from the web based survey and spatial models and GLM analysis were performed.

4.2. Data collection and unity measures

The lack of available secondary sources for the phenomenon under study was overcome by collecting data through a web survey including questions related to working condition, socio-economic status, residential condition, ICTs use, mobility and lifestyle divided in 6 sections, for a total of 86 items (see chapter 2).

Among these variables, for the purpose of spatial analysis[1], only the variable of current geographical location (living place) was considered (var code B02) as an "aggregated individual-level attribute" by count on spatial basis (Haining 2003: 48)

The online questionnaire was the only available way to collect data, in the absence of official statics sources. However the web channel for data collection and the absence of available data from a reference population introduce bias sources such as sampling coverage (see chapter 3).

It is then worth to underline that the lack of a probability sampling and the common relatively low participation rates do not allow any generalization of our findings beyond the participants.

After collecting micro data, the following step had to deal with introducing in the analysis macro territorial data (land use patterns[2]) with reference to accessible housing market, net of available transport infrastructures, density of businesses location and innovation-oriented local markets.

Macro data collection was performed by combining together data from different available sources: social and economic figures, mostly provided by

- 2011 census data from ISTAT (Italian Statistics Bureau)
- Italian Revenue Agency - Real Estate Market Observatory (OMI), 2012
- MIUR data (Education, University and Research Italian Ministry), 2017
- Statistical Atlas from National Research Project (PRIN) "Territori Post-Metropolitani Come Forme Urbane Emergenti: Le Sfide Della Sostenibilità, Abitabilità E Governabilità" (secondary source[3]) .

[1] Being able to locate each respondent on the map would allow to detect for phenomena of spatial segregation within areas which are isolated from the economy and labour market of the city region (by having lower levels of transport and communication connections, universities and firms linkages, research centres, firms' clusters and so on).

[2] "<<Land use>> refers to the spatial distribution of functions such as living, working, shopping and recreating, and determines the relative proximity of different types of activities. spatial distribution of various activities" (Van Aacker, 2010)

[3] The project was founded by Education, University and Research Italian Ministry in 2012 e coordinated by Polytechnic University of Milan. Statistical atlas is available as online resource at www.postmetropoli.it/atlante

Because of different reference data source data were collected during different years: collection time range goes from 2011 to 2017. Nevertheless variations of values from one year to the other have proved to be minimum and do not introduce relevant measurement errors in the analysis.

Variable of interests selected and collected for each area are as described in table 1.

Table 1: High specialized professionals and higher level technicians by gender, age and educational level

FAMILY	**SUB-GROUPS**	**SINGLE VARIABLES**
Business-based indicators	Innovation-driven market	▪ Number of patent applications (2014) ▪ Number of research centres and universities (2017)
	Infrastructures indexes	▪ Index of available telematics and telephonic infrastructures (2012) ▪ Index of available economic infrastructures (2012)
	Local economy	▪ Number of local active companies (2014) ▪ Real estate market: renting in euros per month per sqm 2017
Population.based indicators	Labour market	▪ Unempolyment rate (2011) ▪ High-medium specialized jobs rate (2011)
	Mobility	▪ Work and study-related daily mobility out of municipality (2011)

(Source: own elaboration)

Qualitative and Quantitative indicators were collected for both indexes[4] (Index of available telematics and telephonic infrastructures and Index of available economic infrastructures).
Index of available economic infrastructures was mainly built considering road networks, rail networks, ports and airports facilities , energy and environment ones and postal and banking endowments (for details on Indexes construction see the Appendix).

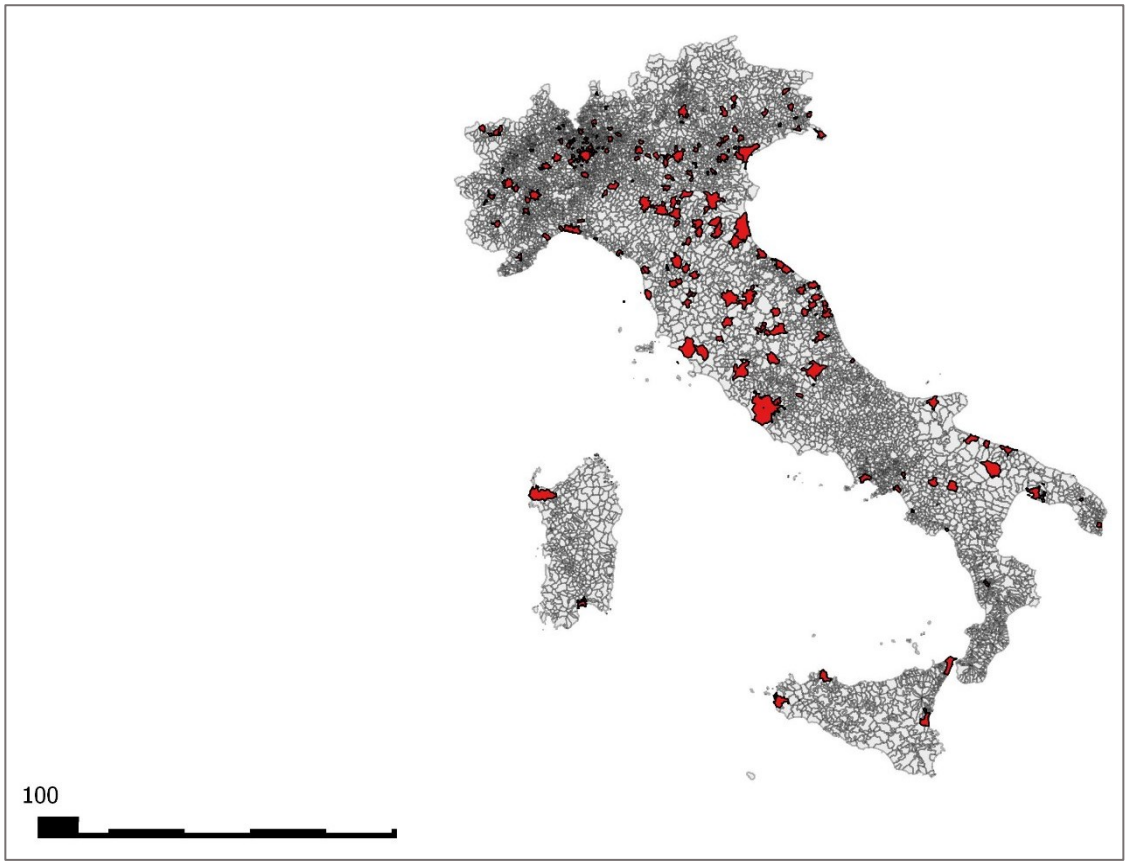
Last point involved describing areal spatial units of analysis as polygons; manipulations were performed on administrative original maps (shapefile format) available from Italian census data 2011 .
We collected 440 answers spread over 208 Italian municipalities.

Total number of Italian municipalities (with reference to last CENSUS data from 2011) is 8094, that is our survey was based just on 2,5% of total municipalities number. These few 208 sparse areas made our

[4] Each index is a ratio between composite indicator of infrastructural endowment and indicator of potential/ real demand. Values are compared to the national mean (=100), so that values above 100 are found for areas with infrastructural endowment higher than national mean, whereas values below 100 are found for areas with infrastructural endowment lower than the national.

national map full of holes due to municipalities without data collected from any respondents (see *graph 1*).

Graph 1: Current living municipalities of respondents (var code B02)

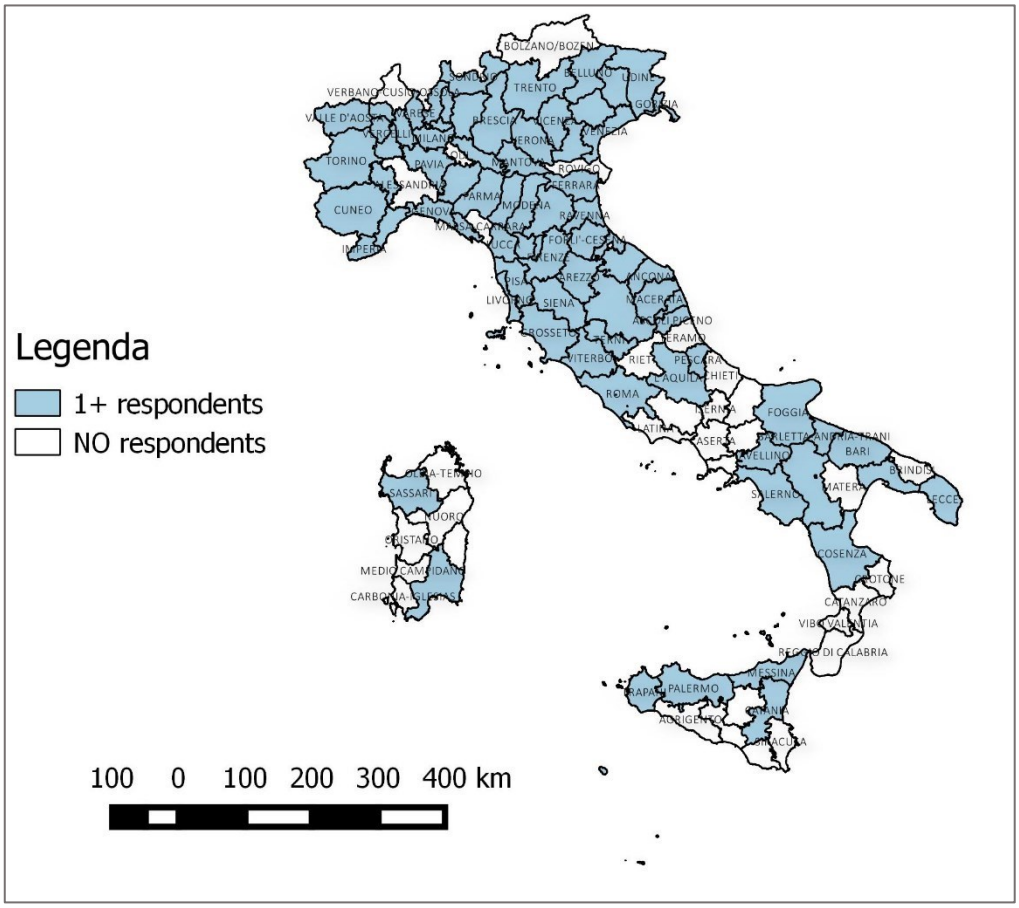


(Source: own elaboration by using Q-GIS software)

For this reason we decided to perform our analysis on aggregated data at province levels (NUTS 3 according Eurostat definition[5]) as shown in *graph 2*.

[5] NUTS classification (Nomenclature of Territorial Units for Statistics) is a hierarchical system for dividing up the economic territory of the EU in order to better allow compared analysis.

Graph 2: Current living provinces of respondents (aggregation for var code B02)

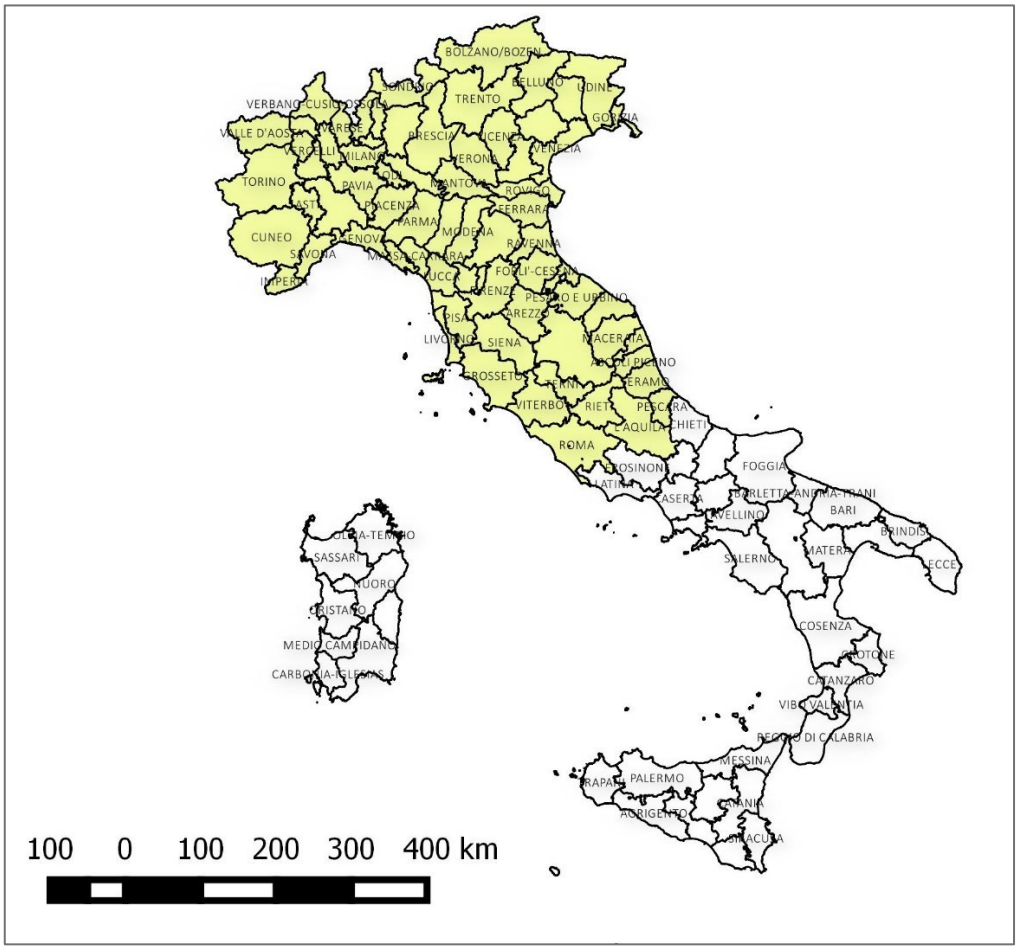


(Source: own elaboration by using Q-GIS software)

As next step, because of high number of missing responses in Southern Italy, we cut narrowed the map from North to centre Italy decreasing the number of province areas with missing values to 9, as from *graph 3*.

As a result our geographical level unity of interests were set to 70 provinces from Piedmont to Abruzzo.

Graph 3: Provinces included in the analysis (aggregation and sub-set for var code B02)



(Source: own elaboration by using Q-GIS software)

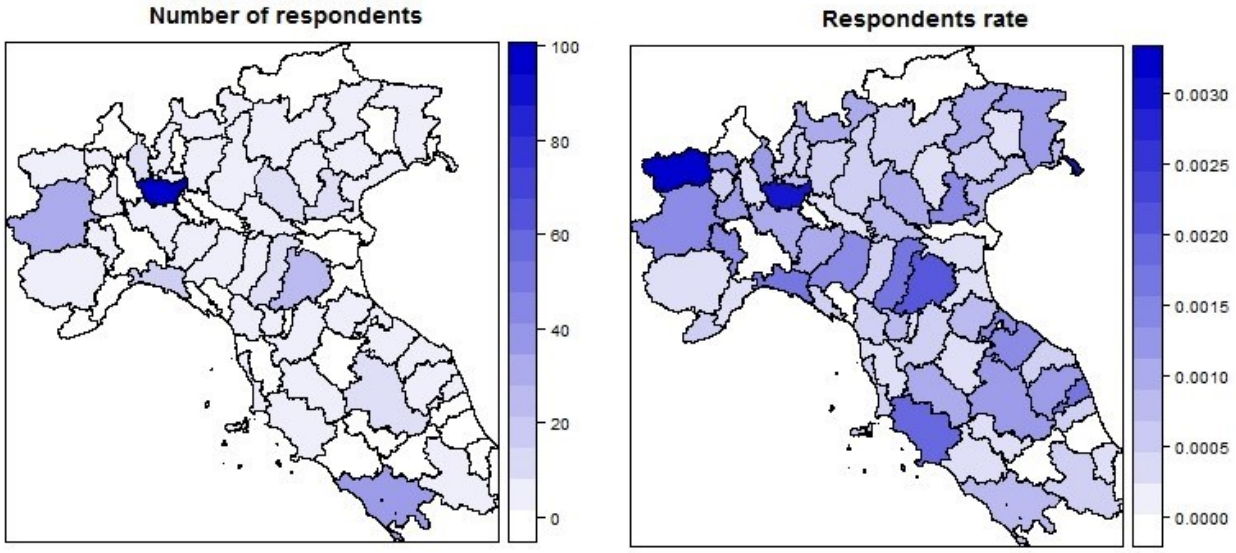
The decision of aggregating data to a larger spatial scale made us loose details and implies addressing the 'Modifiable Areal Unit Problem' (MAUP)[6], associated with the use of data aggregated to geographical areas. The MAUP is related to the fact that the observed aggregated values will vary according to how we draw our area boundaries, by bringing changes in the apparent geographical distribution of the variable in question. Although considering that "changes in aggregation may change outcomes" (Bivand et al. 2013 : 263), we decided to perform data aggregation and selection since this was the only way to be able to perform spatial analysis on lettice data.

Number of respondents (variable RESP_N) was then measured as simple count number and weighted for the resident population (variable RESP_POP) for each of the 70 provinces identified.

Graph 4 display distribution maps both for RESP_N (absolute numbers) and RESP_POP (rates).

[6] The term was first coined by Stan Openshaw (1984: 3) who highlighted that "the areal units (zonal objects) used in many geographical studies are arbitrary, modifiable, and subject to the whims and fancies of whoever is doing, or did, the aggregating".

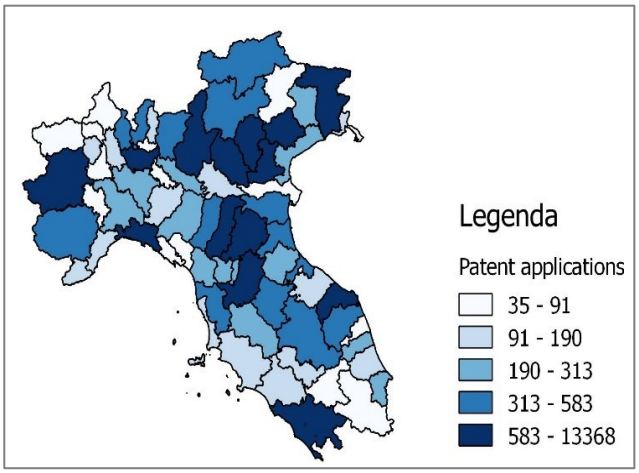
Graph 4: Respondents number VS Respondents rate (count for var code B02)

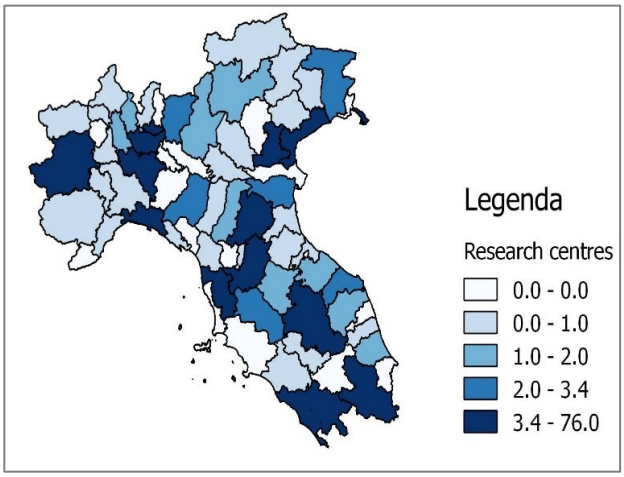


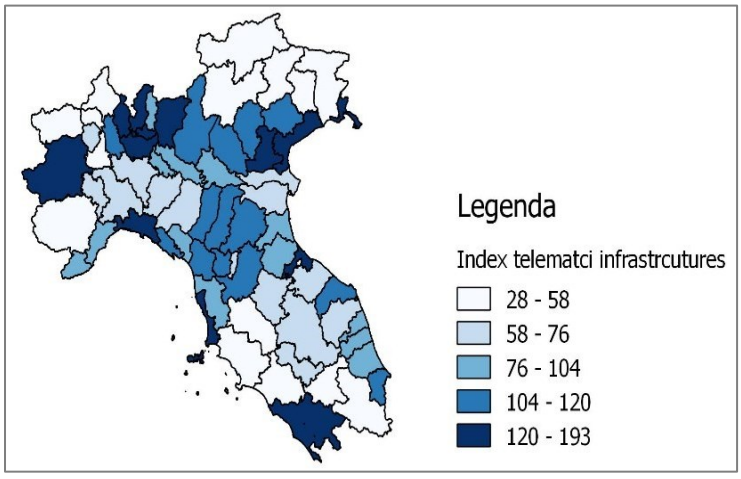
(Source: own elaboration by using Rstudio software)

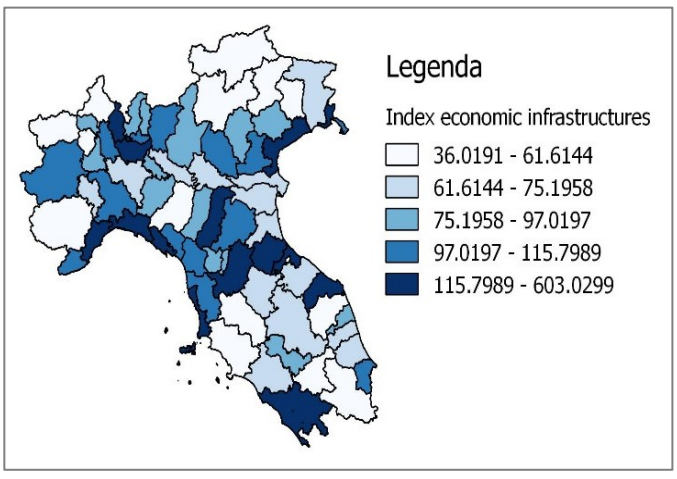
Same distribution maps were printed out for the nine selected macro variables

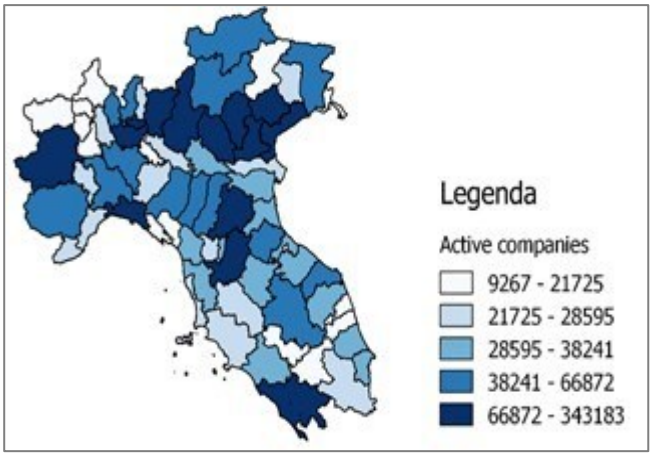
Graph 5: Distribution maps for collected macro variables at province level

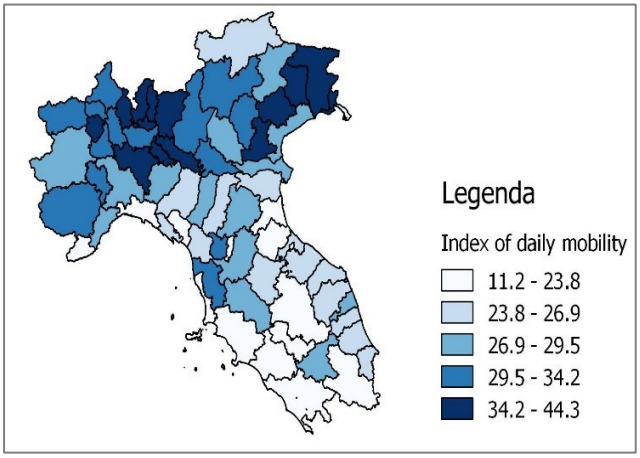


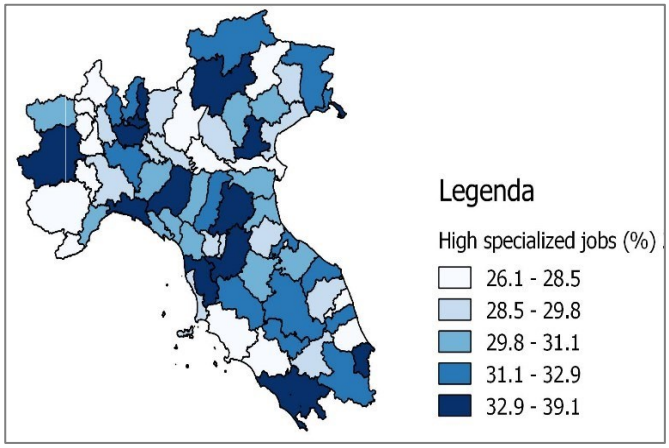


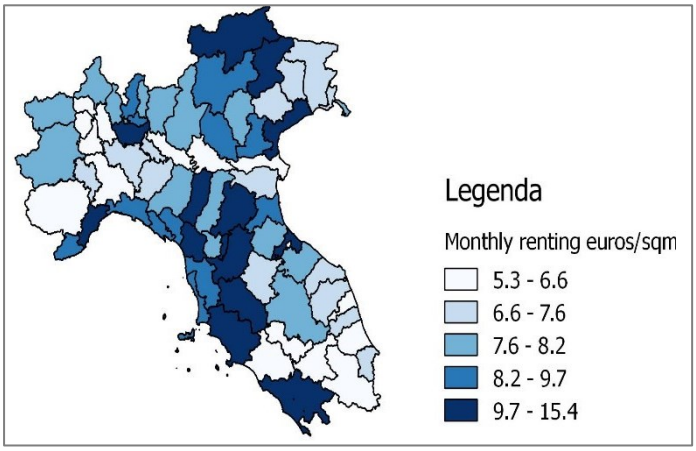


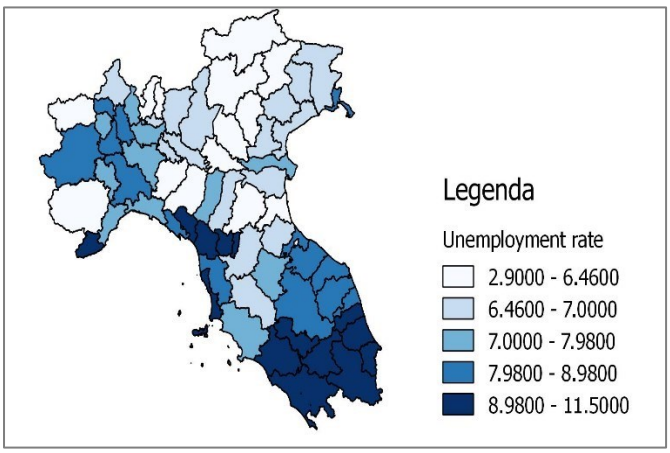












(Source: own elaboration by using Q-Gis software)

These maps allow to easily display socio-economic trends and infrastructural endowments in Northern and Central areas of Italy.

As we may notice the innovation factor (measured by number of patents and research centres) reaches its peaks in central nodes such as Torino Milano, Bologna, Firenze and Roma. Telematics and Economics infrastructures are much more developed in the North, exception made for Roma, Ancona and Pescara. The North-West region were we expected to find “increasing number of advanced infrastructures

connecting the nodes" (see chap 1) does not shows specific connecting trends. On the countrary high scores of infrastructures indexes are detected in the North Est and in the Central areas, along the route which runs from Venezia to Milano – Varese and the one which goes from Genova to Livorno passing by Bologna.

Despite these differences daily mobility index reaches high scores in Northern provinces, with highest scores in municipalities near Milan and Venice, two central areas where people are prone to work/study in, but not to live in due to high prices of the real estate market.

Active companies are numerous in big municipalities such as Torino, Milano, Genova, Bologna, Firenze, Roma; but they do built a real corridor in the North-Est along a route which goes from Venezia to Milano.

The labour market for highly specialized workers does not reveal any strong pattern, but it confirms the importance of central areas (Torino, Genova, Milano, Bologna, Firenze, Roma) as pulling ones, albeit some of them have to cohabitate with high unemployment rates, too (i.e. Torino and Roma).

To summarize, macro territorial data do not seem to support the hypothesis of the polycentric city region model (Scott 2001) as the new central rising area in terms of innovation and leading economic activities, infrastructural endowments and mobility frequencies. Important roles are played, instead, by main provinces (metropolitan areas) which collect higher rates of respondents (*graph 4*) and at the same time display high scores for local macro socio- economic phenomena investigated in *graph 5*.

4.3 **Spatial dependence analysis**

This chapter aims to test the hypothesis of spatial proximity and identify “central areas” where, according to the theory, clusters should form.

Spatial statistics analysis on areal data is usually implemented in economic geography, regional sciences epidemiological studies, but not very much applied in sociological studies yet.

Spatial data take the form of polygon entities defined by boundaries for which observations are available: the observed data are frequently aggregations within the boundaries, such as population counts, and areal entities themselves constitute the units of observation.

Thus every data collected on the field could be mapped and analysed by considering their spatial nature.

We first used **Q GIS** software to built , modify and display maps as polygons with coordinates. Maps can be portrayed with different layers (each one corresponding to a specific scale) and enriched with tables reporting economic data.

This complex file is built of different components and can be esported with a shapefile format into Rstudio statistical software where analysis and models were performed.

4.3.1 Neighbourhood structure and weights matrixes specification

We had to deal with lattice data displayed on a grid and a prerequisite to any analysis is to define neighbourhood relations among areas. A fundamental question is <<how to define ‘neighbours’?>> .
The first step is to define the sets of neighbours for each observation, that is the spatial relationships that exists among areas; the second to assign weights to each neighbour relationship in order to build the so-called “spatial proximity matrix” (Haining 1990; Bailey and Gatrell 1995).

In a strict sense, two spatial units are either neighbors or not and choosing the neighbour criterion to be used it means to define which relationships between observations are to be given a non-zero weight.

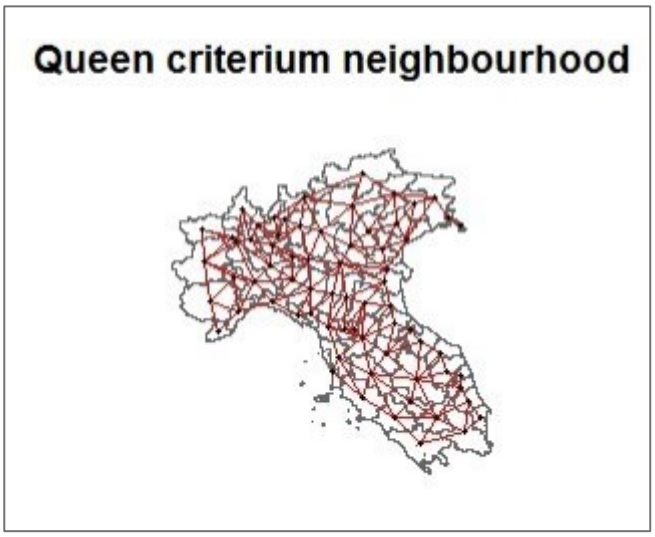
Neighbourhood is defined by a proximity matrix (or weights matrix), W.

In W matrix wij measures proximity between Ai and Aj areas (by convention wii = 0) which could be defined according different criteria.

Here below the 4 criteria chosen and their graph-based representations

1. **Continguity based neighbourhood (QUEEN criterium[7])** force tracts sharing boundary points to be taken as neighbours;

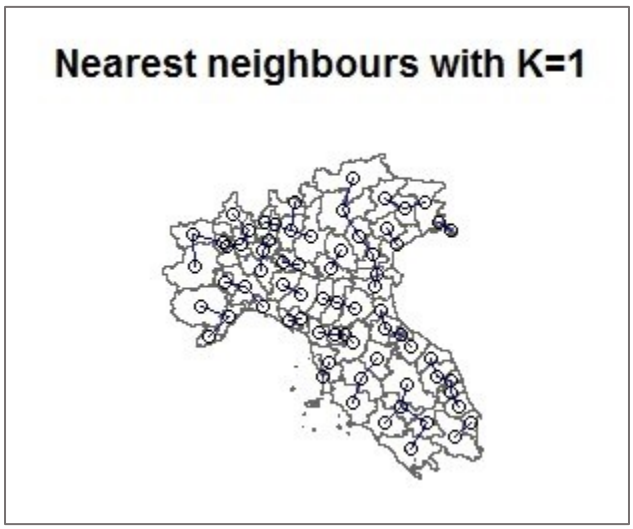
Graph 6: Neighbourhood structures map for the contiguity based criterium display where each graph represents a neighbour relation



(Source: own elaboration by using Rstudio software)

2. **K-nearest neighbours** ensures that all areas have K neighbours. "The K = 1 object is also useful in finding the minimum distance at which all areas have a distance-based neighbour"
(Bivand et al. 2013 : 269)

Graph 7: Neighbourhood structures map for the K=1 criterium display where each graph represents a neighbour relation

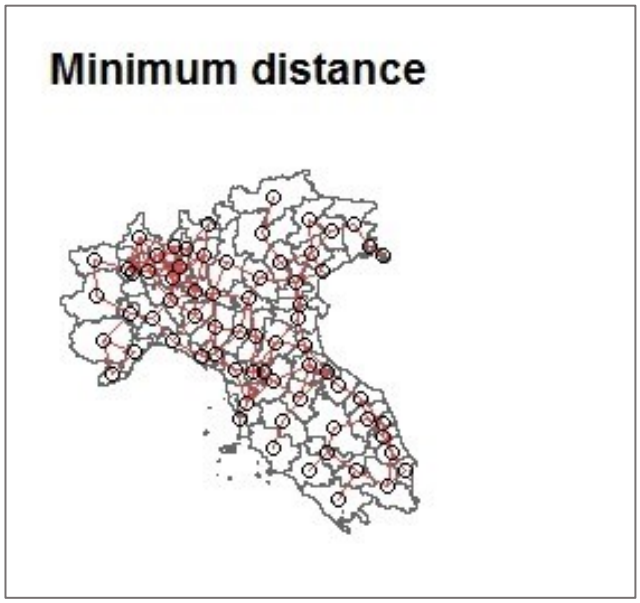


(Source: own elaboration by using Rstudio software)

[7] Queen and Rook are defined by analogy with movements on a chessboard. Where at least one point on the boundary of one polygon is within the snap distance of at least one point of its neighbour: queen criterium. Where at least two boundary points must be within the snap distance of each other: conventional name of a 'rook' relationship (mainly used when the data are known to be arranged in a regular, rectangular grid).

3. **Critical cut-off neighbourhood as distance-based neighbours** defined as points representing areas within distance bands.

Graph 8: Neighbourhood structures map for the minimum distance so that K=1 criterium display where each graph represents a neighbour relation



(Source: own elaboration by using Rstudio software)

The identified neighbour relationships need to be given a value: after the list of sets of neighbours for our study area is established, we proceed to assign spatial weights to each relationship.

Weights wij in the W weights matrix representation will set to zero, where j is not a neighbour of i. The simplest neighborhood definition (criterium 1) is provided by the binary connectivity matrix , whose elements are

$$Wij = \begin{cases} 1 & \text{if regions i and j share a boundary} \\ 0 & \text{otherwise} \end{cases}$$

We may want to expand our idea of a neighborhood to include regions that are close, but not necessarily adjacent (criterium 2).

$$Wij = \begin{cases} 1 & \text{if the centroid of region j is one of the K nearest to the centroid of region i} \\ 0 & \text{otherwise} \end{cases}$$

Alternatively we can define the neighbours by some parametric function of distance (criterium 3, where δ is set as the highest value among the overall minimum distances computed for K=1). For example, if dij is the distance (Euclidean, city-block, or any other distance metric) between the centroids of regions i and j, we could choose

$$Wij = \begin{cases} 1 & \text{if } dij < \\ 0 & \end{cases}$$

Instead of being just binary, weights may be generalised to reflect something that we "know" about the spatial dependency process, for example that it falls with increasing distance or that it increase with the length of borders shared.

We are basically setting continuous-value weights such as length of a shared border between polygons or inverse distance between points. Rather than expressing spatial influence as a binary value (adjacent or not), it is often expressed as a continuous value.

The simplest approach is to use inverse distance (the further away, the lower the value):

$$wij = \begin{cases} 1/dij & \text{if regions i and j share a boundary} \\ 0 & \text{otherwise} \end{cases}$$

Weights were then standardised[8] for each of the 4 matrixes defined, so that row sums in each matrix are unity; this way we can take into account total numbers of neighbours.

[8] Row-standardised weights increase the influence of links from observations with few neighbours, while binary weights vary the influence of observations — those with many neighbours are up-weighted compared to those with few (Bivand R,2007: 25).

4.3.2 Spatial autocorrelation tests

In many cases, it is common to assume that observations are independent and identically distributed, but this may not be the case when working with spatial data. Observations could not be independent in case there exist some correlation between previously defined neighbouring areas.

This phenomenon is what is defined as Spatial Autocorrelation[9]: instead of the classical correlation coefficient measuring the degree of association between 2 different variables, we can measure the spatial correlation as the association among neighbour areas which display different local values of the same variable. In previous chapter we hipotesized the number of Ipros itself to be the most relevant pull factor (or centripetal force) for Ipros to locate in a specific area, so we made an hypothesis of spatial autocorrelation for the variable geographical location.

There exist different tests for identifying spatial autocorrelation; among them we have global and local ones.

Global tests test the whole study area for spatial autocorrelation, assuming the spatial process is the same everywhere.

A global index can suggest clustering but cannot identify individual clusters: this is why we use local indicators of spatial association (LISAs), local versions of well-known global indexes.

"The main purpose of such indicators is to provide a local measure of similarity between each region's associated value and those of nearby regions" (Waller, Gotway 2004: 237).

The most used index for testing global and local spatial correlation is Moran's index (I)[10]: it is calculated as a ratio of the product of the variable of interest (Z) and its spatial lag, with the cross-product of the variable of interest, and adjusted for the spatial weights used (w_{ij}):

$$I = \frac{1}{s^2} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} (Z(s_i) - \bar{Z}) (Z(s_J) - \bar{Z})}{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}}$$

Where $s^2 = \frac{1}{n} \sum_{i=1}^{n} (Z(s_i) - \bar{Z})^2$

[9] Autocorrelation (whether spatial or not) is a measure of similarity (correlation) between nearby observations and rely on Tobler's (1970) first law according to which "the values of a variable in near-by locations are more similar or related than values in locations that are far apart (Fischer, Wang 2011:7).

[10] Moran's Index is widely used in regional studies of several disciplines and it was developed by Patrick Alfred Pierce Moran in 1950 (Moran P.A.P., "Notes on Continuous Stochastic Phenomena". *Biometrika 37 (1)*: 17-23).

"In the definition of Moran's I equation, we assess the spatial similarity of deviations of each regional count Zi with the overall mean regional count Z" (Ibidem :229).

Moran's index (I) expected value is defined as follows

$$\mathbb{E}(I) = -\frac{1}{n-1}$$

So that "when neighbouring regions tend to have similar values (i.e., the pattern is clustered), I will be positive. If neighbouring regions tend to have different values (i.e., the pattern is regular), I will be negative."
(Ibidem: 227)

Whereas when there is no correlation between neighbouring values, the expected value of I approaches zero.

Moran's I reflects a spatially weighted form of Pearson's correlation coefficient (Ibidem, ivi) where null hypothesis (H0) is the absence of spatial autocorrelation[11].

It is extremely important to underline the fact that outcome might vary according to the spatial weights assigned and neighbourhood criterium set; this is why we have set 5 different weights matrixes to be tested.

Furthermore suppose that regions with large population sizes tend to occur near each other: any autocorrelation in the data may simply be due to relationships among the population sizes and not to any spatial pattern in the disease counts. Consequently Waller and Gotway (2004: 229) suggests to "replace regional disease counts with regional crude incidence pro- portions (rates), seeking to remove or at least lessen the impact of population heterogeneity" .

Our variable of interest (RESP_N) was then adjusted for heterogeneous population sizes by dividing each count by the regional population size (RESP_POP). Moran index was then calculated for each different weights matrix previously built (see paragraph 5), by using a permutation test to introduce spatial randomness[12]. This procedure allow to assess the distribution of the autocorrelation index under random

[11] "Inference for a global index of spatial autocorrelation derives from the null distribution (i.e., the distribution of the index under the null hypothesis), so that observed values of the index falling in the tails of this distribution suggest significant spatial autocorrelation" (Waller and Gotway, 2004: 225-226) .

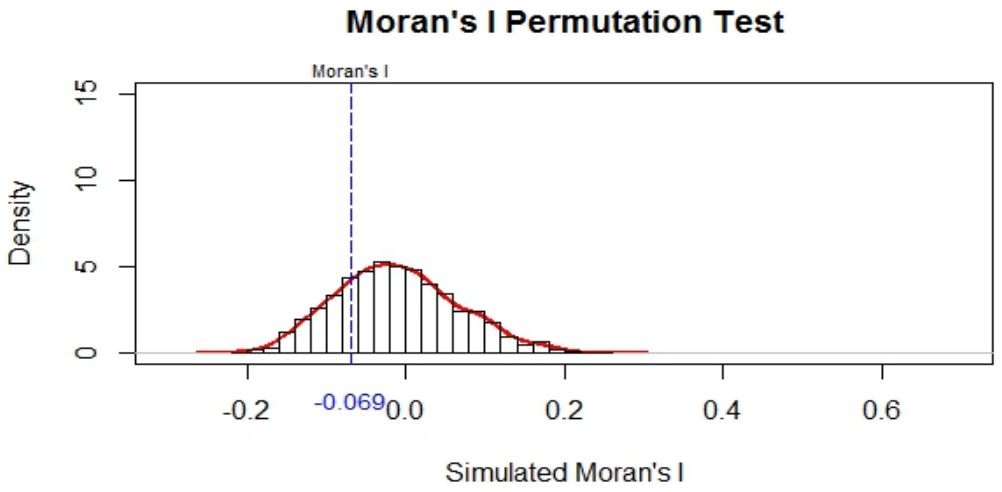
[12] Permutation introduces spatial randomness by simulating H0, so that for every iteration Moran index is calculated and we can display the empirical distribution over to total permutations. The Moran index for our observed value can be displayed against the distribution: if it falls out we can reject H0.

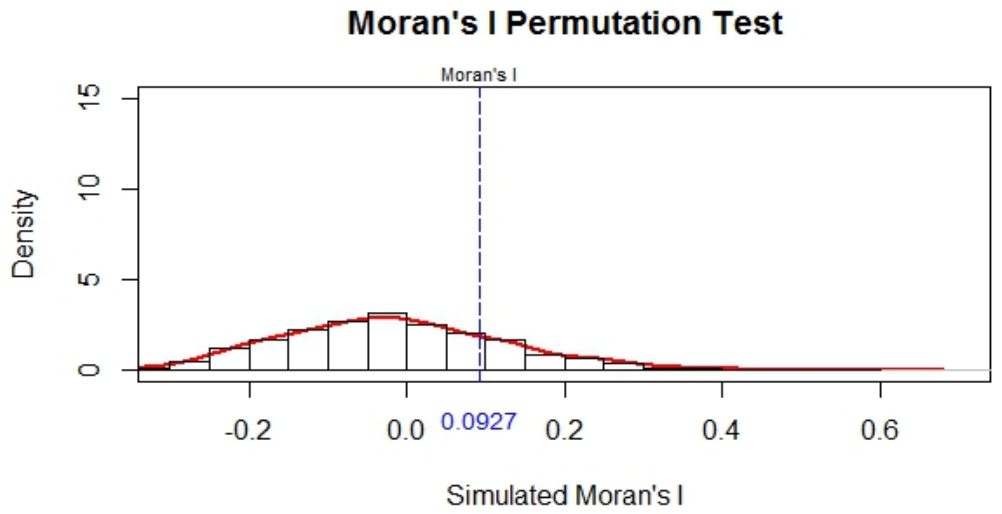
assignment of the values observed to areas: “the values are simply reassigned among the (fixed) locations” (Ivi: 226).

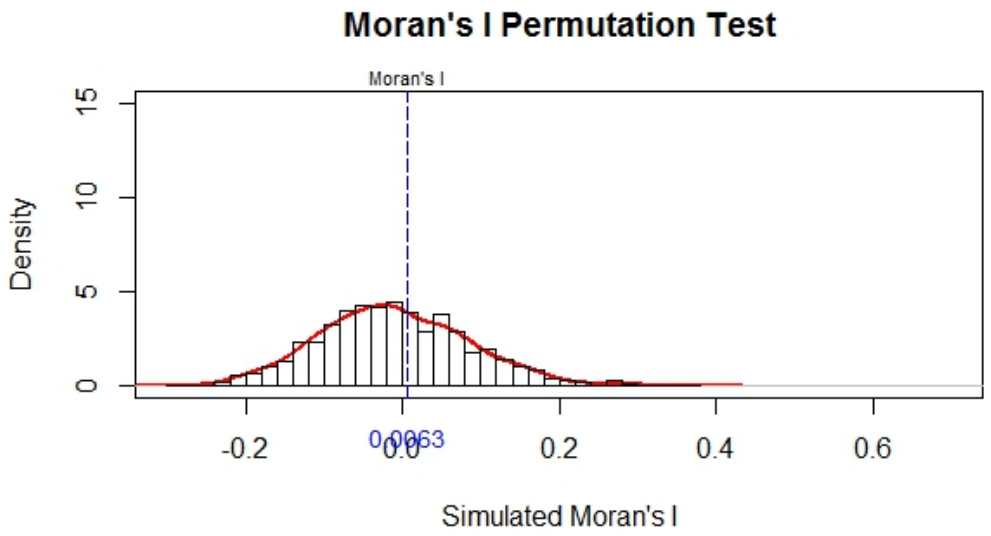
Simulations no. was set as 900 and Moran’s I computed under Normal distribution hypothesis.

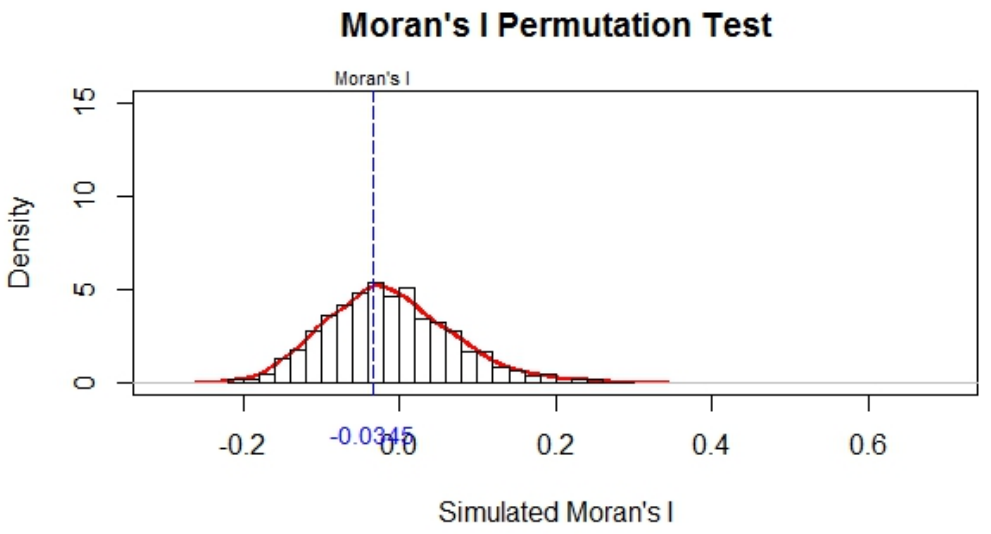
No significant values for Moran’s I were detected (see table 2).

MORAN'S I computation and tested under spatial randomness hypothesis for NORMAL DISTRIBUTION		
WEIGHTS MATRIX	MORAN'S I	Plot: MORAN'S I VS SIMULATED MORAN'S I HISTOGRAM

Binary weights: Queen criterium
 	
`bperm <-`
`moran.mc(Italy_prov$RESP_POP,`
`listw=Italy_prov_nbW, nsim=999,`
`na.action=na.omit)`
	`Monte-Carlo simulation of Moran's I`
	
	`data: Italy_prov$RESP_POP`
	
	`weights: Italy_prov_nbW`
	
	`number of simulations + 1: 1000`
	
	`statistic = -0.0025,`
	`observed rank = 524,`
	`p-value = 0.77`
	`alternative hypothesis:`
	`greater`
		

Binary weights: K-nearest neighbours (K=1)
 	
`bpermK1 <-`
`moran.mc(Italy_prov$RESP_POP,`
`listw=Italy_prov_nb_K1W,`
`nsim=999, na.action=na.omit)`
	`Monte-Carlo simulation of Moran's I`
	
	`data: Italy_prov$RESP_POP`
	
	`weights: Italy_prov_nb_K1W`
	
	`number of simulations + 1: 1000`
	
	`statistic = -0.0927,`
	`observed rank = 788,`
	`p-value = 0.212`
	`alternative hypothesis:`
	`greater`
		

Binary weights: minimum distance for K=1
 	
`bpermdist<-`
`moran.mc(Italy_prov$RESP_POP,`
`listw=dist_NB_1W, nsim=999,`
`na.action=na.omit)`
	`Monte-Carlo simulation of Moran's I`
	
	`data: Italy_prov$RESP_POP`
	
	`weights: dist_NB_1W`
	
	`number of simulations + 1: 1000`
	
	`statistic = 0.0063,`
	`observed rank = 602,`
	`p-value = 0.398`
	`alternative hypothesis:`
	`greater`
		

Generalized weights:
inverse of the distance for Queen criterium
 	
`bpermglistd<-`
`moran.mc(Italy_prov$RESP_POP,`
`listw=Italy_prov_nbglistd,`
`nsim=999, na.action=na.omit)`
	`Monte-Carlo simulation of Moran's I`
	
	`data: Italy_prov$RESP_POP`
	
	`weights: Italy_prov_nbglistd`
	
	`number of simulations + 1: 1000`
	
	`statistic = -0.0345,`
	`observed rank = 413,`
	`p-value = 0.587`
	`alternative hypothesis:`
	`greater`
		

(Source: own elaboration by using RStudio software)

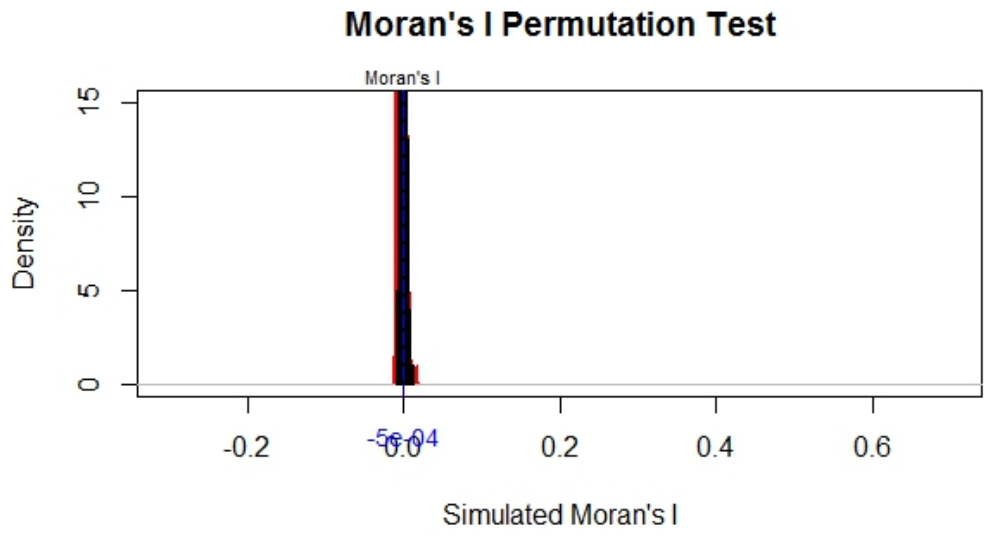
The same permutation test was then performed by assuming our count variable RESP_POP to follow a Poisson distribution.

In other words Monte Carlo test (still based on 999 simulations) under the constant risk hypothesis allows to generate independent Poisson counts with mean rn i (the overall incidence proportion times the population size) for each area S. Basically "we simulate counts based on these expectations but apply Moran's I as defined in previous equation, comparing each regional count to the overall mean regional count" (Waller, Gotway 2004: 231).

Anyway spatial dependence was hard to detect even under the constant risk hypothesis (see table 3).

Table 3: Moran's Index values computed for Poisson distribution with 4 different adjacency matrixes

MORAN'S I computation and tested under spatial randomness hypothesis for POISSON counts		
WEIGHTS MATRIX	MORAN'S I	Plot: MORAN'S I VS SIMULATED MORAN'S I HISTOGRAM

Binary weights: Queen criterium
 	
`bpermEBI<-`
`EBImoran.mc(Italy_provnona$RESP_N`
`, Italy_prov$POP, Italy_prov_nbW,`
`nsim=999)`
	`Monte-Carlo simulation of`
	`Empirical Bayes Index`
	
	`data: cases:`
	`Italy_provnona$RESP_N,`
	`risk population: Italy_prov$POP`
	`weights: Italy_prov_nbW`
	`number of simulations + 1: 1000`
	
	`statistic = -5e-04,`
	`observed rank = 708,`
	`p-value = 0.292`
	`alternative hypothesis: greater`
		

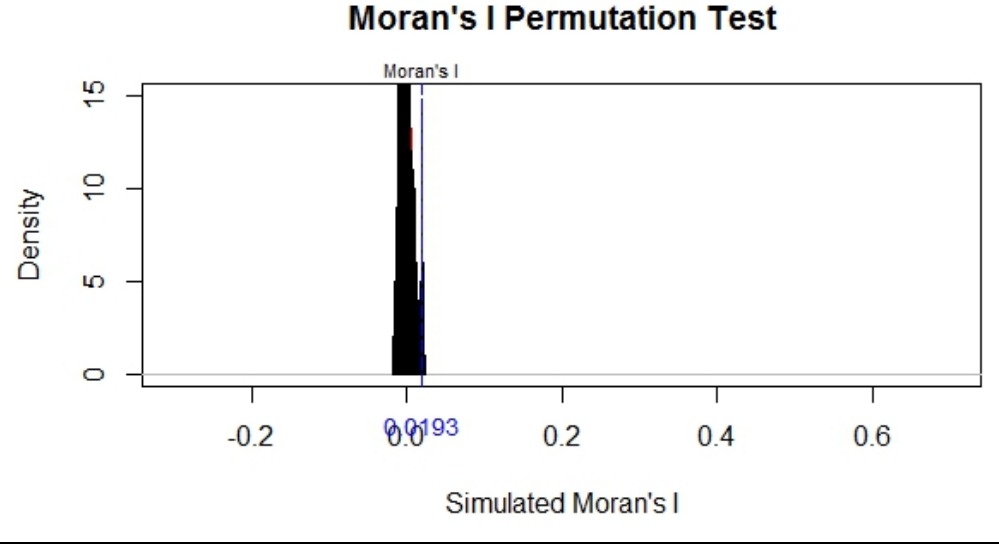
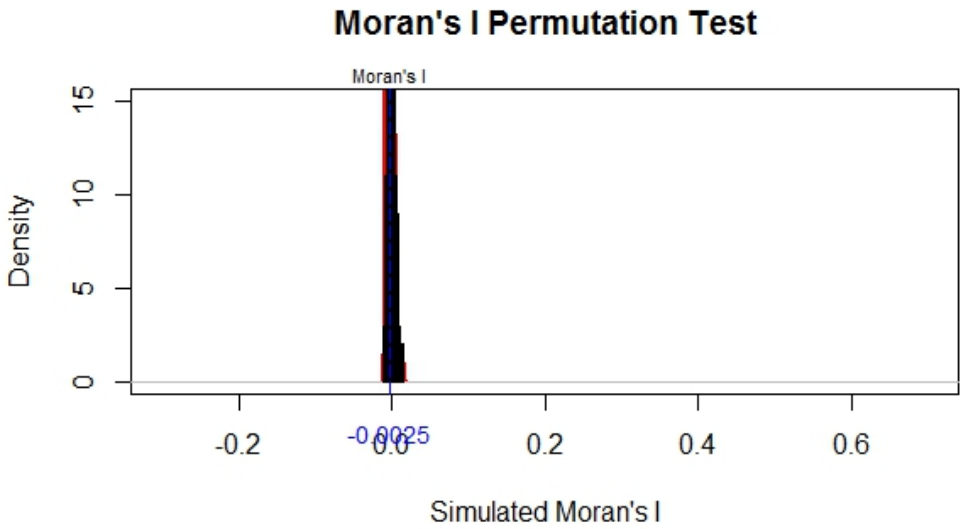
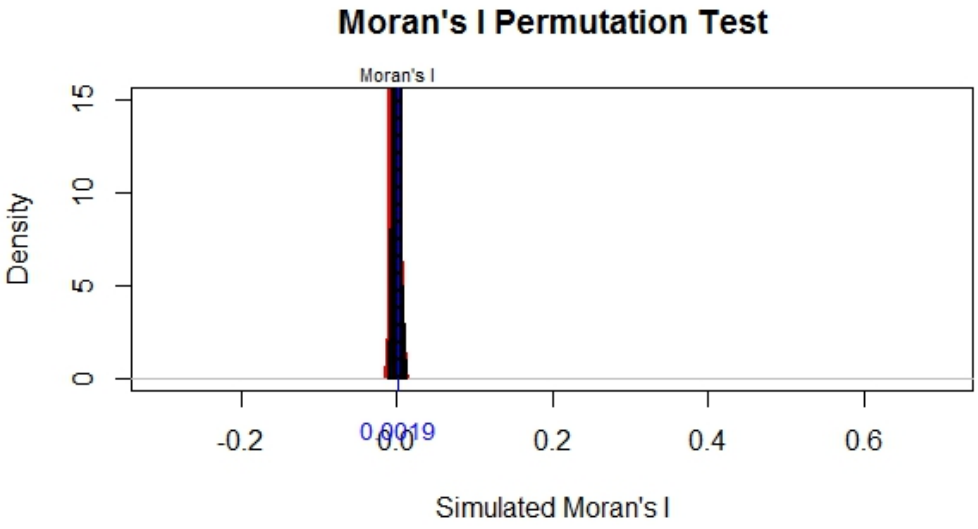
Binary weights: K-nearest neighbours (K=1)
 	
`bpermEBIK1<-`
`EBImoran.mc(Italy_provnona$RESP_N`
`, Italy_prov$POP,`
`Italy_prov_nb_K1W, nsim=999)`
	`Monte-Carlo simulation of`
	`Empirical Bayes Index`
	
	`data: cases:`
	`Italy_provnona$RESP_N,`
	`risk population: Italy_prov$POP`
	`weights: Italy_prov_nb_K1W`
	`number of simulations + 1: 1000`
	
	`statistic = 0.0193,`
	`observed rank = 978,`
	`p-value = 0.022`
	`alternative hypothesis: greater`
		

Table 3: Moran's Index values computed for Poisson distribution with 4 different adjacency matrixes

Binary weights: minimum distance for K=1
 	
`bpermEBIdist<-EBImoran.mc(Italy_provnona$RESP_N, Italy_prov$POP, dist_NB_1W, nsim=999)`
	`Monte-Carlo simulation of Empirical Bayes Index`
	
	`data: cases: Italy_provnona$RESP_N,`
	`risk population: Italy_prov$POP`
	`weights: dist_NB_1W`
	`number of simulations + 1: 1000`
	`statistic = 0.0063,`
	`observed rank = 602,`
	`p-value = 0.476`
	`alternative hypothesis: greater`
		

Generalized weights:
inverse of the distance for Queen criterium
 	
`bpermEBIglistd<-EBImoran.mc(Italy_provnona$RESP_N,Italy_prov$POP,Italy_prov_nbglistd, nsim=999)`
	`Monte-Carlo simulation of Empirical Bayes Index`
	
	`data: cases: Italy_provnona$RESP_N,`
	`risk population: Italy_prov$POP`
	`weights: Italy_prov_nbglistd`
	`number of simulations + 1: 1000`
	`statistic = 0.0019,`
	`observed rank = 876,`
	`p-value = 0.124`
	`alternative hypothesis: greater`
	
	`number of simulations + 1: 1000`
		

(Source: own elaboration by using RStudio software)

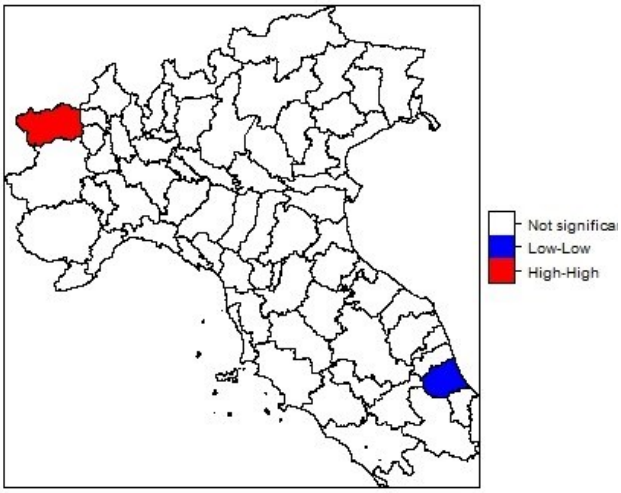
Despite the absence of global autocorrelation for the variable under study, we have also tried to test and plot Moran Local Indicator of Local Association (LISA).

$$\text{Ii} = \sum_{j=1}^{n} w_{ij} sim_{ij} \frac{(Z(s_j) - \overline{Z})}{s^2} \sum_{j=1}^{n} w_{ij} (Z(s_j) - \overline{Z})$$

This Local index is computed for each area S_i and we are returned its value and significance level (p-value), so that we can easily check for clusters of areas with high or low significant correlations.

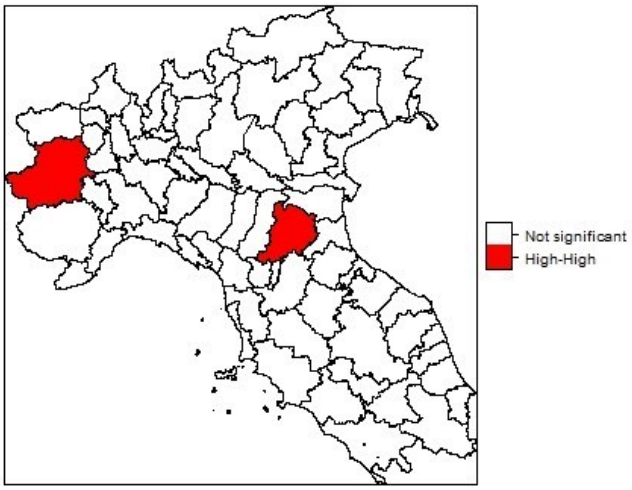
Graph 9: Local Moran Index: significant values and direction of relationship for 4 different adjacency matrixes

Binary weights: Contiguity-based



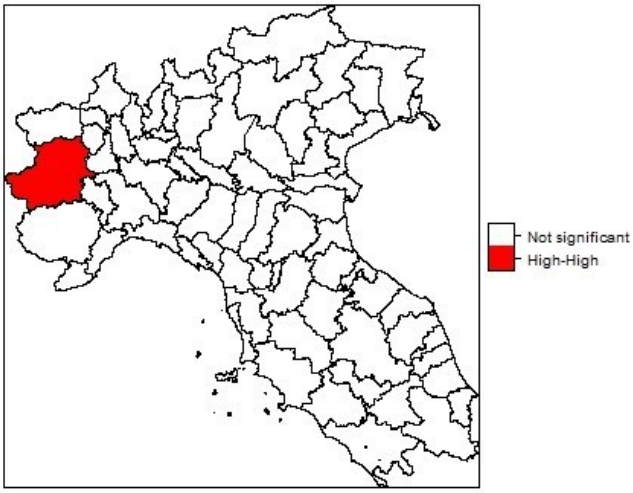
Aosta, Teramo

Binary weights: : K-nearest neighbours



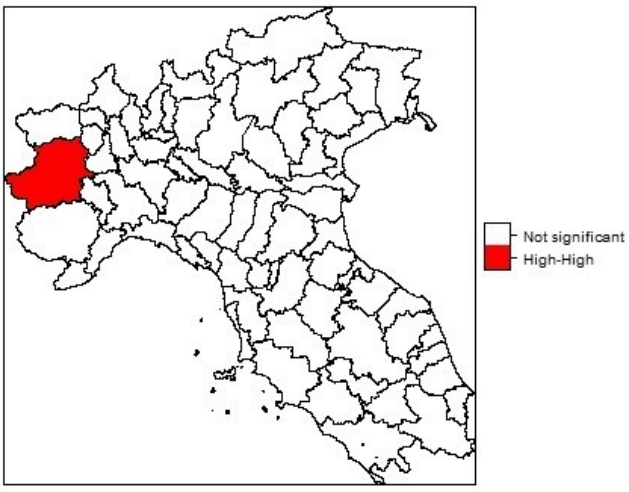
Torino, Bologna

Binary weights: minimum distance for K=1



Torino

Generalized weights: inverse of the distance for contiguity-based



Torino

(Source: own elaboration by using Rstudio software)

Results do not show any specific clustering trend, being coherent with global Moran's index analysis.

However few areas show significant local correlation according to different adiajency matrixes imputed.

Specifically setting contiguity-based criterium positive spatial dependence is found in Aosta (the more Ipros in neighbour areas, the more in Aosta), and negative ones in Teramo (the less Ipros in neighbour areas, the less in Teramo).

Other 3 criteria show just positive dependence especially for Torino and Bologna.

Despite cluster detection failed, we can still consider the relevance of some areas, Torino especially, for Ipros to tend to live near to each other.

A possible explanation to this phenomenon is investigated in paragraph 4.4. by using a GLM.

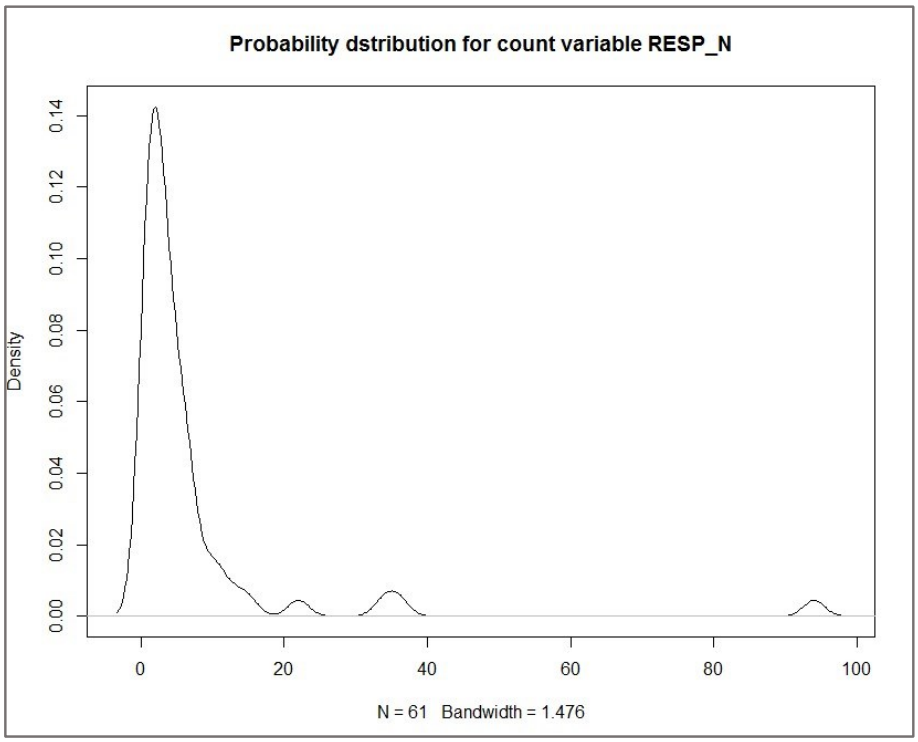
4.4 Generalized linear model for Poisson distribution

Why do Ipros tend to live in some places and not in others? What are main macro-variables influencing the choice? In order to answer this questions we have performed a General Linear regression (GLM)[1] for Poisson family[2].

We assumed underlying Poisson probability distribution for our dependent count variable: number of respondents (RESP_N).

Probability distribution is plot in g*raph 10*

Graph 10: Probability distribution plot for the dependent count variable RESP_N



(Source: own elaboration by using Rstudio software)

In order to better define the models, before listed dependent macro variables were tested for multicollinearity[3].

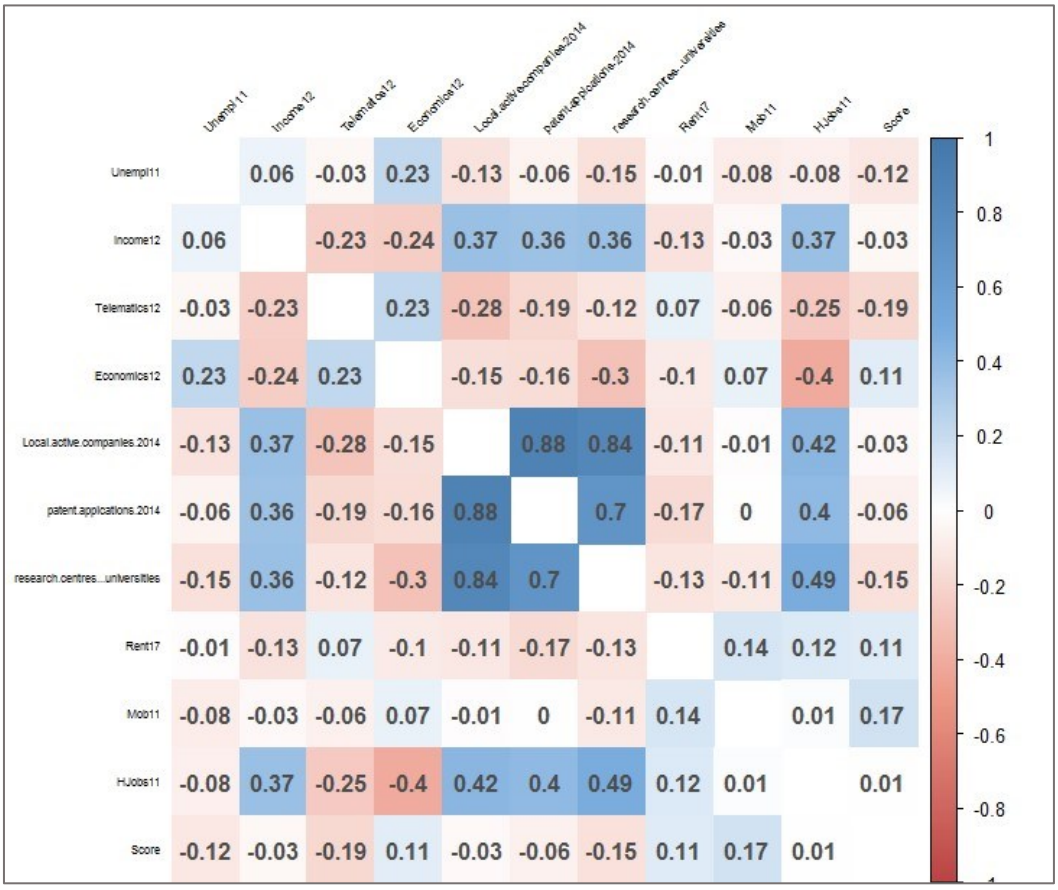
To easily display correlations coefficients a plot was printed in *graph 11*.

[1] Count outcome variables may be log-transformed and analyzed using OLS regression, too. However many issues would arise with this approach, including loss of data due to undefined values generated by taking the log of zero (which is undefined) and biased estimates.

[2] Poisson regression models, mostly used in econometrics and epidemiology literature, are models for describing count data that assume integer values corresponding to the number of events occurring in a given interval.

[3] Multicollinearity increases the standard errors of the estimated coefficients for the variables that are entangled in it.

Graph 11: Correlation coefficient for dependent macro variables



(Source: own elaboration by using RStudio software)

Three main correlations were detected (darker blue cells), all statistically significant with α=0,05 for H0; see *table 4*.

Table 4: Main significant correlations above 0.5

Correlation coefficient	Local.active.companies.2014 (code: ACTIVECOMPANIES_N)	Research.centres. Universities (code: RESEARCH_N)	patent.applications.2014 (cpde: PATENTS_N)
Local.active.companies.2014 (code: ACTIVECOMPANIES_N)		0,84**	0,88**
Research.centres.Universities (code: RESEARCH_N)			0,7**
patent.applications.2014 (cpde: PATENTS_N)			

(Source: own elaboration)

The three mentioned couples of variables proved to be correlated by checking for their Variance Inflation Factor (VIF), too.

VIF[4] quantifies the severity of multicollinearity in an ordinary least squares regression analysis and it measures how much the variance of an estimated regression coefficient is increased because of collinearity. High values are usually set in the range of 5-10 and a VIF of 5 or greater indicates a reason to be concerned about multicollinearity.

During model specification interactions terms were then included for each couple of variables.

At last before including the regressors in the model, in order to get more readable outputs, standardization was performed by substracting the mean (process also known as centering the variables[5]). Our model specification had this formula:

$$RESP\_N = ß_0 + ß\,SPECJOBS\_perc + ß_2\,REALESTATE + ß_3\,MOBILITY\_perc + ß_4\,UNEMPL\_rate + ß_5\,INCOME\_percapita + ß_6\,TELEMATICS\_index + ß_7\,ECONOMICS\_index + ß_8\,ACTIVECOMPANIES\_N + ß_9\,RESEARCH\_N + ß_{10}\,PATENTS\_N + ß_{11}\,(ACTIVECOMPANIES\_N * PATENTS\_N) + ß_{12}\,(ACTIVECOMPANIES\_N * RESEARCH\_N) + ß_{13}\,(PATENTS\_N * RESEARCH\_N) + \varepsilon$$

A stepwise regression method[6] was applied: in each step, a variable is considered for addition to or subtraction from the set of explanatory variables based on The Akaike Information Criterion (AIC)[7]. This way we were returned 5 different models, model V being the most parsimonious with 5/9 regressors kept.

Outputs with regression coefficient estimates are printed in *table 5*.

[4] A VIF for a single explanatory variable is obtained using the r-squared value of the regression of that variable against all other explanatory variables: $VIF_X = \frac{1}{1-R_x^2}$

[5] The main benefit relies on the fact that the interpretation of the coefficients does not change: if you subtract the mean, each coefficient continues to estimate the change in the mean response per unit increase in X when all other predictors are held constant.

[6] In other words, it is a step by step automatic procedure (available for almost every statistical softwares) from base model to more complex one, with lower AIC.

[7] The Akaike Information Criterion (AIC) as well as Bayes Information Criterion (BIC) is based on Maximum Likelihood logarithm. It provides a method for assessing the quality of a model through comparison of related models: It's based on the Deviance, but give penalties for making the model more complicated. Much like adjusted R-squared, its intent is to prevent from including irrelevant predictors.

Table 5: Coefficient estimates over 5 models computed for Poisson regression

GLM for RESP_N (number of respondents)

	Mod I	Mod II	Mod III	Mod IV	Mod V
Rate of medium-high specialised jobs	0.87717***	0.46322***	0.44207***	0.60884***	0.47449***
Number of Active local companies		0.33912***	0.70653***	1.03581***	1.12373***
Number of research centres			-0.38115***	-0.56798***	-0.57758***
Number of patent applications				-0.21013**	-0.32386***
Average income per capita					0.27174**
AIC	430.94	329.03	269.16	260.62	252.96

***p<0.01, ** p<0.05, * p<0.1

(Source: own elaboration by using Rstudio software)

In accordance with our expectations, positive coefficients are shown for rate of medium-high specialised jobs (var code SPECJOBS_perc), number of Active local companies (var code ACTIVECOMPANIES_N) and average income per capita (var code INCOME_percapita): Ipros number would increase in areas where there are higher values of people working in highly specialized sectors, higher numbers of active companies and higher rates of per capita income.

On the contrary research centres (var code RESEARCH_N) and patent applications numbers (var code PATENTS_N) seem to have negative impact on Ipros settling. Further literature should be investigated in order to better explain this relationship.

Nonetheless we could reasonably hypothesize Ipros not to be found close to area with higher research centres and start-up companies since this category of workers mainly work for companies but they are not hired from the companies themselves, neither from research centres. Furthermore a second hypothesis could be pointing out the choice of Ipros for locating not too close to higher competitive counterparts (i.e. start-up companies) on the innovation market.

Excluded regressors which did not prove to be significantly relevant for variance explanation are unemployment rate (var code UNEMPL_rate), telematics index (var code TELEMATICS_index), economics index (var code ECONOMICS_index), real estate market (var code REALESTATE), mobility index (var code MOBILITY_perc), and interaction effects before introduced.

Generally speaking real estate market prices seem not to affect Ipros' decision for living in a place or another (against our hypothesis for the housing market to be the main constraint push factor) and, consequently, daily mobility rate [1] does not prove to be a relevant factor, either.

It is worth to deepen the focus on the relation between income and real estate market: beyond larger units of provinces, the Anova analysis confirms that family income (var C08_02) and real estate market prices by single municipality (var real_estate)[2] are not related one to each-other (i.e. we would expect families with lower income level to prefer living in municipalities with lower real estate markes prices).

[1] This rate was not available as incidence over working population, but only over the whole resident population up to age 64: thus results might have been biased by higher rates of elder people living in a specific area.

[2] In order to perform this analysis a new variable was entered at municipal level on the basis of info gathered by PRIN project "TERRITORI POST-METROPOLITANI COME FORME URBANE EMERGENTI: LE SFIDE DELLA SOSTENIBILITÀ, ABITABILITÀ E GOVERNABILITÀ" (www.postmetropoli.it/atlante), funded by Italian Ministry of Research and Education . This research provide data from OMI 2012 expressed as renting cost in euros per month per squared meter.

Analysis of Variance Table

```
Response: datasetnoNA$real_estate

                          Df  Sum Sq Mean Sq F value Pr(>F)
datasetnoNA$C08_02class    4   18.94  4.7346   0.603 0.6607
Residuals                357 2803.24  7.8522
```

In order to better disentangle this relations, we would better consider other relevant anagraphic variables shaping Ipros' sample profile under observation: as first mentioned in chapter XXX the choice of living in an area compared to another one (and consequently to face higher cost of living or, alternatively, undergo higher commuting experiences) is very likely to be related to anagraphic variables such as age, gender, familiar ties.

4.5 Conclusions and limits of the analysis

This chapter aimed at analysing spatial behaviour of our sample at macro level, by detecting any form of auto-correlation for our dependent variable and testing correlation with local economic structural factors.

Specifically we were interested in addressing two questions:

1. Does the response of neighbours has a direct influence on the response at a particular location ?

2. What are main effects of local macro-economic factors on our response variable?

Question 1. was dealt with by building four different spatial structures marked by specific neighbouring criteria and weights matrixes.

Spatial dependence was tested by computing Moran's index and contrasting it with a spatial random effect model: this method has the advantage of allocating the overdispersion to spatial and non-spatial components, recognizing the inherently spatial nature of the data.

However no spatial dependence was detected and no clusters were found. Only few isolated cases proved to have a significant local auto-correlation when testing for the local Moran's index.

A deep knowledge of the phenomenon under study could suggest to define neighbourhood going beyond pure spatial criteria; it could be appropriate to enter interaction criteria describing social or economic flows between one area and the others (i.e number of journeys, trade data)[3].

In order to better understand these results we have to consider that scale and aggregation levels are two key concepts in spatial data analysis that can be tricky to deal with (Haining 2003:43-44), as well as arbitrary decisions for neighbourhoods spatial relations setting (Ivi: 80-81).
Different phenomena have different spatial and temporal scales at which they operate[4]: processes that operate over a larger extent can be studied at a larger resolution whereas processes that operate over a smaller extent may need to be studied at a more granular level (Levin, 1992).

Thus we might have misrepresent the real scale at which our phenomenon should be studied.

According to Hamnett thesis (1991) highest number of Ipros were indeed found in large metropolitan areas such as Milan, Rome and Torino: clusterization might have been detected, then, at a lower scale on the single metropolitan area. In conclusion the smaller the regions get, the better the real pattern is captured.

However this lower level of analysis was not supported by available collected data which did not cover all areas from a specific domain[5]: not relying on census data we were able to collect data at granular level (zip-codes), but unfortunately high number of areas with missing data forced us to aggregate data at province level.

The way the borders of the areas are drawn (how large, what shape, where) does strongly affect the patterns we see and the outcome of data analysis and since we might have mis-specified the scale of analysis we could not detect any spatial clusters.

This issue is usually referred to as the "Modifiable Areal Unit Problem" (MAUP), whereas the problem of analyzing aggregated data (quite common in the social sciences) is referred to as "Ecological Inference" (Wrigley et al. 1996: 25). Smoothing geographical variation will indeed result in variance reduction and lost of details.

Question 2 was assessed by running a general linear regression for Poisson family distribution on standardized data. AIC was used as the main goodness of fit statistics.

[3] Interaction criteria allow to introduce general weights matrixes (Bavaud 1998), different from the case of binary ones (as previously showed for the case of the inverse of the distance criterium).
[4] Areas used to aggregate data are always arbitrary and should be related to the phenomenon under investigation.
[5] Either analyzing municipality level on a national domain or scaling to zip-codes areas on a metropolitan level, not enough data were available and the grid would have had too many areas with missing response data.

The advantage of using stepwise method were assessed and 5 models were estimated under parsimony, the last one being the best among others.

Against expected results (all included factors should have an impact on our dependent variable), we detected no existing significant relation, net of other regressors, between number of Ipros in a specific area and telematics local infrastructure as well as number of Ipros in a specific area and daily mobility rate for local resident population.

We argue different explanations could be addressed, considering the nature of the available dependent variables (i.e. see note 19) and the lack of micro anagraphic data in the model (gender, age etc.).

To conclude we have to remark the fact that data modelling is especially problematic when dealing with web-suvey collected data: despite efforts for post-stratification and weights adjustment we can hardly make any inference from sample to the population.

These specific results we got provide some insights only for that self-selected sample which answered our online questionnaire.

References

Anselin L. (1992), 'Spatial data analysis with GIS: An introduction to application in the social sciences', technical report, Open Access Publications from the University of California Santa Barbara, National Center for Geographic Information and Analysis (NCGIA)

Anselin, L., Bera, A. K., Florax, R., and Yoon, M. J. (1996). Simple diagnostic tests for spatial dependence. Regional Science and Urban Economics, 26:77–104.

Anselin L., (2005), Spatial Regression Analysis in R A Workbook, Spatial Analysis Laboratory Department of Agricultural and Consumer Economics University of Illinois, Urbana-Champaign Urbana, IL 61801

Anselin, L., Syabri, I., and Kho, Y. (2006). GeoDa: An introduction to spatial data analysis. Geographical Analysis, 38: 5 -22.

Bivand R.S., Pebesma E., Gómez-Rubio V. (2013) Modelling Areal Data. In: Applied Spatial Data Analysis with R. Use R!, vol 10. Springer, New York

Bavaud F. (1998), Models for spatial weights: a systematic look, *Geographical analysis (30)*: 153-171.

Cappellin R., Ferlaino F, Rizzi P. (2012) Introduzione: la città nell'economia della conoscenza. In: Cappellin R., Ferlaino F, Rizzi P. (a cura di) La città nell'economia della conoscenza . AISRe. Franco Angeli (7-27).

Hamnett C. (1991), 'The blind men and the elephant: the explanation of gentrification', Transactions of the Institute of British Geographers 16 (2): 173-189

Haining R. (1990), Spatial Data Analysis in the Social and Environmental Sciences, Cambridge University Press

Haining R. (2003), Spatial data analysis: theory and practice, Cambridge University Press.

Kauffman V.(2011), Rethinking the city. Urban dynamics and motility, Oxford, Routledge (EPFL Press)
Manfred M. Fischer, Jinfeng Wang (2011), Spatial Data Analysis_ Models, Methods and Techniques, SpringerBriefs in Regional Science

Openshaw S. (1983). The modifiable areal unit problem. Norwick: Geo Books

O'Sullivan, D. and Unwin, D. J. (2003). Geographical Information Analysis. Wiley, Hoboken, NJ.

Sassen S. (2007), Una sociologia della globalizzazione , Einaudi.

Simon A. Levin (1992) The problem of pattern and scale in ecology, *Ecology (73)* 6: 1943-1968

Soja W. Edward (2000), Postmetropolis. Critical Studies of Cities and Regions, Oxford, Blackwell

Trevor C. Bailey, Anthony C. Gatrell (1995), Interactive spatial data analysis, Longman Scientific & Technical
Waller, L. A. and Gotway, C. A. (2004), *Applied Spatial Statistics for Public Health Data*, John Wiley & Sons, Hoboken, NJ.

Van Aacker V. (2010), *Spatial and social variations in travel behaviour. Incorporating lifestyles and attitudes into travel behaviour-land use interaction research*, doctoral thesis University of Gent and Technical University of Delft

Wrigley, N., & Lowe, M. (1996), Retailing, consumption and capital: towards the new retail geography. Harlow, GB: Longman.

Zakariya Y. Algamal, (2012),Diagnostic In Poisson Regression Models, *Electronic Journal of Applied Statistical Analysis EJASA 5 (2)*:178-186

Appendix – The Indexes

Index composition for ***Index of Available Economic Infrastructures (2012)***

1. ROAD NETWORK

INDICATORS	DATA SOURCES
Highway network length	AISCAT (Associazione Italiana Società Concessionarie Autostrade e Trafori)
National roads network length	ACI (Automobile Club Italia)
County roads network length	Ministry of Infrastructures and Transports
Secondary roads network length	Ministry of Infrastructures and Transports
Number of toll booths	AISCAT (Associazione Italiana Società Concessionarie Autostrade e Trafori)
Number of motorway service stations	AISCAT (Associazione Italiana Società Concessionarie Autostrade e Trafori)
Number of toll booths with Viacard	AISCAT (Associazione Italiana Società Concessionarie Autostrade e Trafori)
Number of toll booths with Telepass	AISCAT (Associazione Italiana Società Concessionarie Autostrade e Trafori)
County roads maintenance expenditure	Ministry of Infrastructures and Transports

2. RAILWAY NETWORK

INDICATORS	DATA SOURCES
Total railway network length	RFI/FC (Italian Railway System)
Single track railway network length	RFI/FC (Italian Railway System)
Twin track railway network length	RFI/FC (Italian Railway System)
Total railway network length for commercial use	RFI (Italian Railway System)
Number of departing /incoming Eurostars	TRENITALIA
Number of ETR 500 trains	TRENITALIA
Number of ETR 460 trains	TRENITALIA
Number of ETR 450 trains	TRENITALIA
Days when Eurostars pass by	TRENITALIA

3. PORTS

INDICATORS	DATA SOURCES
Number of approaches from seaward	Ministry of Infrastructures and Transports
Approaches length	Ministry of Infrastructures and Transports
Number of approaches for passengers	Ministry of Infrastructures and Transports
Number of oil approaches	Ministry of Infrastructures and Transports
Number of approaches for container goods	Ministry of Infrastructures and Transports
Loading docks size	Ministry of Infrastructures and Transports
Number of approaches for liquid / dry/ pallet goods	Ministry of Infrastructures and Transports
Number of approaches for fishermen	Ministry of Infrastructures and Transports
Number of approaches for pleasure purposes	Ministry of Infrastructures and Transports
Number of approaches for service vehicles	Ministry of Infrastructures and Transports
Cold stores capacity	Ministry of Infrastructures and Transports
Other stores capacity	Ministry of Infrastructures and Transports
Silos capacity	Ministry of Infrastructures and Transports
Number of apporacehs in need for restauration	Ministry of Infrastructures and Transports

4. AIRPORTS

INDICATORS	DATA SOURCES
Landing area size	ENAC (Italian Civil Aviation Authority)
Parking airplanes area size	ENAC (Italian Civil Aviation Authority)
Airstrips length	ENAC (Italian Civil Aviation Authority)
Airstrips width	ENAC (Italian Civil Aviation Authority)
Airstrips surface	ENAC (Italian Civil Aviation Authority)
Public expenditures	Ministry of Infrastructures and Transports
ENAV expenditures	Ministry of Infrastructures and Transports
ENAV incomes	Ministry of Infrastructures and Transports
Subjects	Ministry of Infrastructures and Transports
Commercial facilities, passengers area	Ministry of Infrastructures and Transports
Commercial facilities airside area	Ministry of Infrastructures and Transports
Commercial facilities in landside area	Ministry of Infrastructures and Transports
Distance from the municipality	ENAC (Italian Civil Aviation Authority)

5. ENERGY-ENVIRONMENT NETWORKS

INDICATORS	DATA SOURCES
Water-purifying plants	ISTAT
Water distribution system capacity	ISTAT
Tanks capacity	ISTAT
Gas-served land surface	SNAM (Italian natural gas infrastructure company)
Gas-served inhabitants	SNAM (Italian natural gas infrastructure company)
Gas distribution network length	SNAM (Italian natural gas infrastructure company)
Total electricity power consumption	GRTN (National Transmission Network Operator; electric power industry)
Net production of electricity power	GRTN (National Transmission Network Operator; electric power industry)
Total production of urban wastes	APAT (Agency for Environmental Protection and Technical Services)
Total production of special wastes	APAT (Agency for Environmental Protection and Technical Services)
Total supplied water	ISTAT
Feeders media pressione	SNAM (Italian natural gas infrastructure company)
Rete a bassa pressione	SNAM (Italian natural gas infrastructure company)
Total selective waste collection	APAT (Agency for Environmental Protection and Technical Services)
Urban selective waste collection	APAT (Agency for Environmental Protection and Technical Services)
Recycled special wastes (in terms of material and power)	APAT (Agency for Environmental Protection and Technical Services)

6. BANKING AND OTHER SERVICES NETWORKS

INDICATORS	DATA SOURCES
Number of Postal offices	Poste Italiane (National postal services provider)
Number of Banking desks	National Bank of Italy
Add. Consulenza per installazione di elaboratori elettronici	ISTAT (ASIA)
Consultans for software and informatics equipments management	ISTAT (ASIA, Statistical archive active for companies)
Workers in charge for Electronic Data processing	ISTAT (ASIA, Statistical archive active for companies)
Workers in charge for electronic machines maintenance	ISTAT (ASIA, Statistical archive active for companies)

Workers in charge for accounting, corporate consultancy, tax consultancy, auditing activities	ISTAT (ASIA, Statistical archive active for companies)
Workers in charge for administrative consultancy	ISTAT (ASIA, Statistical archive active for companies)
Workers in charge for testing and technical products analysis	ISTAT (ASIA, Statistical archive active for companies)
Number of postal offices with faster products shipments	Poste Italiane (National postal services provider)
Number of postal offices with fax	Poste Italiane (National postal services provider)
Number of postal offices with postamat	Poste Italiane (National postal services provider)
Number of total active Pos terminals	Banca d'Italia (National postal services provider)
Number of available ATM terminals	Banca d'Italia (National postal services provider)

Index composition for of ***Available Telematics and Telephonic Infrastructures (2012)***

1. TELEMATICS AND TELPHONIC INFRASTRCUTURES	
INDICATORS	**DATA SOURCES**
Resident population with ADSL coverage higher than 50%	Osservatorio Banda Larga (Broadband Internet Observatory)
Resident population in municipalities covered by HDSL	Osservatorio Banda Larga (Broadband Internet Observatory)
Resident population in municipalities covered by UMTS	Osservatorio Banda Larga (Broadband Internet Observatory)
Resident population with ADSL coverage higher than 76%	Osservatorio Banda Larga (Broadband Internet Observatory)

Appendix – The Questionnaire

Stili di vita per freelancers e liberi professionisti

L'Università degli Studi di Milano e l'Università del Piemonte Orientale la invitano a concedere parte del suo prezioso tempo per arricchire con la sua testimonianza lo studio delle nuove forme di lavoro autonomo

In particolare stiamo conducendo una INCHIESTA su 6 dimensioni della vita quotidiana dei LAVORATORI FREELANCERS e PROFESSIONISTI (lavoro, luogo di residenza, status socio-economico, uso delle tecnologie, mobilità sul territorio, tempo libero).

Tutte le domande sono numerate in sequenza e completare il questionario richiederà circa 15 minuti.

Le risposte saranno trattate in forma anonima: i dati ottenuti saranno resi noti solo a livello di sintesi statistica e utilizzati per fini di ricerca scientifica.

Per ringraziarvi della collaborazione saranno sorteggiati tra i partecipanti

- 1 carta regalo della Feltrinelli del valore di 100

euro - 4 carte regalo della Feltrinelli del valore di

50 euro l'una.

Benvenuto!

Leggere attentamente le istruzioni sottostanti prima di procedere con la compilazione:

Istruzioni:

1. Il questionario è rivolto a tutti quei lavoratori freelancer o professionisti che traggono dal proprio reddito da lavoro la principale fonte di sostentamento. Nel caso di secondo lavoro, quindi, è importante ai fini della rilevazione che la maggior fonte di guadagno provenga dall'attività di freelancer. In caso contrario vi chiediamo di non rispondere al questionario.

2. E' possibile scegliere più di una opzione solo dove espressamente indicato

3. Quando si sceglie l'opzione "altro" indicarne sempre il contenuto

Ci sono 66 domande all'interno di questa indagine.

Sezione A Condizione lavorativa precedente

[A01]Quando hai iniziato a lavorare come freelancer? *

mese

anno

Es. se si è iniziato nel **MARZO 2013** scrivere **03** per il mese e **2013** per l'anno

[A02]Prima di diventare freelancer eri

*

Scegli **solo una** delle seguenti:

- ○ studente/ssa
- ○ lavoratore/trice dipendente
- ○ titolare di impresa con lavoratori alle dipendenze
- ○ stagista o lavoratore/trice con contratto atipico (CO.CO.CO, CO.CO.PRO)
- ○ inoccupato (condizione di chi non ha mai lavorato)
- ○ Altro [ ]

[A021]Per quale motivo hai lasciato il precedente lavoro e sei diventato freelancer? *

Scegli **solo una** delle seguenti:

- ○ ho perso il lavoro / è finito lo stage / non mi hanno rinnovato il contratto
- ○ prospettive di carriera
- ○ sono andato/a in pensione
- ○ motivi familiari
- ○ Altro [ ]

Sezione A.1 Condizione lavorativa attuale

[A03]Dai un nome alla tua professione (anche in inglese se vuoi) *

Scrivere la propria risposta qui:

[A04]Come definiresti lo status legato al tuo lavoro ? *

Scegli **solo una** delle seguenti:

- ○ lavoratore/trice autonomo/a con partita iva non iscritto ad un ordine
- ○ lavoratore/trice indipendente con contratto atipico (CO.CO.CO, CO.CO.PRO)
- ○ libero professionista con partita iva iscritto ad un ordine
- ○ ricercatore/trice, dottorando/a o borsista
- ○ Altro

[A041]Indica il tuo regime fiscale *

Scegli **solo una** delle seguenti:

- ○ Regime dei minimi
- ○ Regime forfettario
- ○ Regime ordinario

[A05]La tua cassa contributiva è *

Scegli **solo una** delle seguenti:

- ○ la gestione separata INPS
- ○ cassa specifica di un ordine professionale
- ○ artigiani e commercianti INPS
- ○ nessuna
- ○ Altro

[A051]Godi di una previdenza integrativa privata (assicurazioni, fondi pensione etc.)? *

Scegli **solo una** delle seguenti:

- ○ Sì
- ○ No

[A06]Nel tuo lavoro sei principalmente *

Scegli **solo una** delle seguenti:

- ○ pagato a prestazione / progetto
- ○ pagato a tempo

[A07]Qual è il settore della tua professione?

*

Scegli **solo una** delle seguenti:

- ○ consulenza / servizi alla persona
- ○ commercio
- ○ artiginato
- ○ produzione industriale
- ○ consulenza /servizi specialistici alle imprese
- ○ Altro

[A071]Specifica l'ambito dei servizi erogati *

Scegli **solo una** delle seguenti:

- ○ informatica e comunicazione
- ○ finanza e assicurazioni
- ○ indagini di mercato, ricerca
- ○ formazione, life-long learning
- ○ arte e intrattenimento
- ○ Altro

se la tua professionalità tocca diversi ambiti indica quello nel quale ti riconosci di più

[A08]Di solito dove lavori? *

Scegli **solo una** delle seguenti:

- ○ a casa
- ○ in un co-working
- ○ presso il cliente
- ○ nel mio studio
- ○ Altro

Sezione A.2 Informazioni sui clienti / committenti

[A09]Quanti sono i clienti per i quali lavori attualmente?

*

Scrivere la propria risposta qui:

[A10]I tuoi clienti sono principalmente *

Scegli **solo una** delle seguenti:

- ○ imprese
- ○ no profit/ privato sociale/ cooperative
- ○ singoli privati
- ○ pubblica amministrazione
- ○ Altro

[A11]Lavori più spesso *

Scegli **solo una** delle seguenti:

- ○ in completa autonomia
- ○ in collaborazione con altri freelancers
- ○ in collaborazione con clienti / personale dell'ente
- ○ Altro

[A12]Come hai conosciuto la tua rete di collaboratori? *

Scegliere **tutte** le corrispondenti:

- ☐ passaparola / amicizie comuni
- ☐ tramite ACTA
- ☐ avete frequentato lo stesso percorso di studi
- ☐ frequentate lo stesso coworking
- ☐ precedenti esperienze di lavoro
- ☐ Altro:

[A13]Quanto spesso lavori in loco presso i tuoi clienti ? *

Scegliere la risposta appropriata per ciascun elemento:

quotidianamente	più di una volta a settimana	una volta a settimana	più di una volta al mese	una volta al mese	meno di una volta al mese	mai
○	○	○	○	○	○	○

[A14]Rispetto al luogo in cui vivi i la maggioranza dei tuoi clienti si trova *

Scegli **solo una** delle seguenti:

- ○ nello stesso quartiere
- ○ nella stessa città
- ○ nella stessa area metropolitana (capoluogo + comuni limitrofi)
- ○ nella stessa provincia
- ○ nella stessa regione
- ○ nelle regioni limitrofe
- ○ in altre regioni italiane
- ○ all'estero

Sezione B. Condizioni abitative

[B01]Dove vivevi prima di diventare un freelancer? *

Comune

CAP

[B02]Dove vivi attualmente ? *

Comune

CAP

(indicare anche se il domicilio è rimasto lo stesso indicato nella domanda precedente B01)

[B03]Vivi in casa con la tua famiglia di origine ? *

Scegli **solo una** delle seguenti:

- ○ Sì
- ○ No

Con "famiglia di origine" si intendano madre, padre, fratelli, sorelle, nonni, nonne.

[B031]Quando ti sei trasferito nel tuo attuale domicilio? *

mese

anno

[B04]Se potessi scegliere senza vincoli economici e familiari, vorresti vivere in un altro luogo in Italia? *

Scegli **solo una** delle seguenti:

- ○ Sì
- ○ No

Con vincoli economici si intendano i fattori legati al costo della vita e degli affitti, con vincoli familiari si intendano fattori legati alla vicinanza al proprio nucleo familiare

[B041]In quale altro comune vorresti vivere in Italia? *

Scrivere la propria risposta qui:

[B042]Qual è il principale motivo per cui vorresti vivere in questo luogo ? *

Scegli **solo una** delle seguenti:

- ○ c'è una buona rete dei trasporti (es. stazioni metro, treni ad alta velocità, piste ciclabili)
- ○ c'è una buona rete per le comunicazioni (es. diffusione wi-fi, fibra ottica)
- ○ è facile l'accesso a servizi utili al mio lavoro
- ○ ci sono aria pulita, spazi verdi, tranquillità
- ○ è minore il degrado urbano e il rischio di criminalità
- ○ ci sono potenziali clienti concentrati sul territorio
- ○ l'ambiente culturale è affine al mio
- ○ Altro [ ]

Il riferimento è al luogo indicato nella domanda precedente (B041)

Sezione C. Condizione socio-economica

[C01]Indica il tuo genere *

Scegli **solo una** delle seguenti:

- ○ Femmina
- ○ Maschio

[C02]Indica la tua età *

Scrivere la propria risposta qui:

[C03]Indica la tua cittadinanza *

Scegli **solo una** delle seguenti:

- ○ italiana dalla nascita
- ○ italiana acquisita
- ○ straniera europea
- ○ straniera non europea
- ○ apolide

[C04]Qual è il tuo titolo di studio più alto? *

Scegli **solo una** delle seguenti:

- ○ licenza elementare
- ○ licenza media
- ○ qualifica professionale
- ○ diploma di scuola superiore
- ○ laurea triennale
- ○ master di I livello
- ○ laurea magistrale / vecchio ordinamento
- ○ master di II livello
- ○ dottorato di ricerca / specializzazione post lauream
- ○ Altro

[C05]Qual è il titolo di studio dei tuoi genitori? *

Scegliere la risposta appropriata per ciascun elemento:

	licenza elementare	licenza media	qualifica professionale	diploma di scuola superiore	laurea triennale	master di I livello	laurea magistrale / vecchio ordinamento	master di II livello	dottorato di ricerca / specializzazione post lauream
Padre	○	○	○	○	○	○	○	○	○
Madre	○	○	○	○	○	○	○	○	○

Sezione C.1 Condizioni del nucleo familiare

[C07]Qual' è la tua condizione coniugale? *

Scegli **solo una** delle seguenti:

- ○ single
- ○ convivente
- ○ sposato/a
- ○ divorziato/a
- ○ vedovo/a

Nella definzionione di "single" rientrano anche quelle persone che hanno una relazione ma non convivono

[C08]Indica il numero dei familiari che convivono con te *

Scrivere la propria risposta qui:

Nel caso si viva da soli (famiglia mononucleare) indicare "0"

[C081]Quanti sono i familiari che vivono con te e che non hanno un reddito da lavoro o da pensione? *

Scrivere la propria risposta qui:

[C082]Chi contribuisce in misura maggiore al reddito familiare? *

Scegli **solo una** delle seguenti:

- ○ io
- ○ il/la mio/a partner
- ○ Altro

[C083]Esistono strategie familiari per fronteggiare la tua instabilità / discontinuità lavorativa? *

Scegli **solo una** delle seguenti:

- ○ Sì
- ○ No

Si intenda strumenti di compensazione del reddito familiare in grado di sopperire a tuoi eventuali periodi di mancanza di lavoro. Es. redditi provenienti da rendite fondiarie / catastali, reddito "sicuro" del partner lavoratore dipendente, alto reddito del partner etc.

[C084]In quale fascia si colloca il vostro reddito familiare annuale lordo? *

Scegli **solo una** delle seguenti:

- ○ meno di 8.000 euro
- ○ 8.000 -12.000 euro
- ○ 12.000 -18.000 euro
- ○ 18.000 -25.000 euro
- ○ 25.000 -35.000 euro
- ○ 35.000 -50.000 euro
- ○ 50.000 -70.000 euro
- ○ 70.000 euro e oltre

[C09]Oltre al reddito da lavoro come freelancer hai altre fonti di reddito ? *

Scegliere **tutte** le corrispondenti:

- ☐ no
- ☐ si svolgo un secondo lavoro
- ☐ si percepisco la pensione
- ☐ si, ho altri patrimoni
- ☐ Altro: _____

[C10]In quale fascia si colloca il tuo reddito annuale da lavoro lordo? *

Scegli **solo una** delle seguenti:

- ○ meno di 5.000 euro
- ○ 5.000 -8.000 euro
- ○ 8.000 -12.000 euro
- ○ 12.000 -18.000 euro
- ○ 18.000 -25.000 euro
- ○ 25.000 -35.000 euro
- ○ 35.000 euro e oltre

Sezione D. Uso delle tecnologie

[D01]Quale di questi dispositivi tecnologici possiedi? *

Scegliere **tutte** le corrispondenti:

- ☐ laptop / notebook (computer portatile)
- ☐ tablet
- ☐ Personal Computer (dispositivo fisso)
- ☐ telefono fisso
- ☐ telefono cellulare senza connessione internet
- ☐ telefono smartphone / blackberry
- ☐ Altro: _____

[D02]Quali reti usi per lavorare? *

Scegliere **tutte** le corrispondenti:

- ☐ rete internet domestica con cavo
- ☐ rete internet domestica wifi
- ☐ rete wifi pubblica
- ☐ rete internet per smartphone
- ☐ rete voce per telefono cellulare o smartphone
- ☐ rete voce per telefono fisso
- ☐ Altro: _____

[D03]In generale con quale frequenza comunichi con i tuoi clienti ? *

Scegliere la risposta appropriata per ciascun elemento:

	Tutti i giorni	Più di una volta a settimana	Una volta a settimana	Più di una volta al mese	Una volta al mese	Meno di una volta al mese	Mai
di persona	○	○	○	○	○	○	○
per telefono	○	○	○	○	○	○	○
su Skype	○	○	○	○	○	○	○
per e-mail	○	○	○	○	○	○	○
Per sms (oppure whatsapp)	○	○	○	○	○	○	○
Tramite social network	○	○	○	○	○	○	○

Nel caso di più clienti prova a fare una stima generale

Sezione E. Mobilità e spostamenti sul territorio

[E01]Nell'ultimo mese quanto spesso ti sei spostato/a sul territorio per motivi di lavoro *

Scegliere la risposta appropriata per ciascun elemento:

	Più volte al giorno	Almeno una volta al giorno	Più di una volta alla settimana	Una volta alla settimana	Più di una volta al mese	Una volta al mese	Meno di una volta al mese	Mai
Nel tuo quartiere	○	○	○	○	○	○	○	○
In altre aree della tua città	○	○	○	○	○	○	○	○
Nell'area metropolitana (capoluogo + comuni limitrofi)	○	○	○	○	○	○	○	○
Nella tua regione di residenza	○	○	○	○	○	○	○	○
Nelle regioni limitrofe	○	○	○	○	○	○	○	○
In altre regioni italiane	○	○	○	○	○	○	○	○
All'estero	○	○	○	○	○	○	○	○

Es. incontro clienti / riunioni / team-work

[E02]Nell'ultimo mese quanto spesso ti sei spostato/a sul territorio per accedere a servizi utili al tuo lavoro? *

Scegliere la risposta appropriata per ciascun elemento:

	Più volte al giorno	Almeno una volta al giorno	Più di una volta alla settimana	Una volta alla settimana	Più di una volta al mese	Una volta al mese	Meno di una volta al mese	Mai
Nel tuo quartiere	○	○	○	○	○	○	○	○
In altre aree della tua città	○	○	○	○	○	○	○	○
Nell'area metropolitana (capoluogo + comuni limitrofi)	○	○	○	○	○	○	○	○
Nella tua regione di residenza	○	○	○	○	○	○	○	○
Nelle regioni limitrofe	○	○	○	○	○	○	○	○
In altre regioni italiane	○	○	○	○	○	○	○	○
All'estero	○	○	○	○	○	○	○	○

Es. servizi manutenzione, rifornimento presso negozi, servizi manutenzione hardware-software, attività di risotrazione, coworking, formazione, fiere etc.

[E03]Nel tuo ultimo giorno lavorativo ti sei spostato dalla tua abitazione per motivi di lavoro *

Scegli **solo una** delle seguenti:

○ Sì

○ No

Considerare abitazione il luogo dove si dimora, anche temporanemante

[E031]Quali sono gli spostamenti che hai percorso nel tuo ultimo giorno lavorativo? *

Scrivere la propria risposta qui:

Si intenda per spostamento il tragitto percorso tra due tappe.

Se non si sono effettuati spostamenti indicare il luogo in cui si è rimasti durante l'arco dell'intera giornata

Indica la sequenza usando " -" come nell'esempio. *Es. (casa)- panetteria - lavoro - scuola - casa : 4 spostamenti*

[E032]Quanti di questi spostamenti (con riferimento alla domanda E03) sono legati a ad attività NON lavorative? Indicare il numero *

Acquisti e spesa	
Motivi familiari (far visita a parenti, prendere i bambini a scuola etc.)	
Tempo libero (hobby, sport, intrattenimento)	
associazionismo, volontariato	

[E033]Quanti di questi spostamenti (con riferimento alla domanda E03) sono stati per motivi di lavoro? Indicare il numero *

Scrivere la propria risposta qui:

Con "motivi di lavoro" si intenda qualsiasi spostamento legato all'esercizio (anche indiretto) dell'attività lavorativa: riunione id lavoro, incontri con clienti, accesso ad una postazione di lavoro (coworking, studio, ufficio), accesso a servizi postali, di manutenzion e consulenza, rivendite etc.

Es. Se ci si è recati in posta per inviare una raccomandata ad un cliente allora bisognerà considerare il tragitto percorso uno **spostamento per motivi di lavoro**

Sezione E.1 Mobilità e mezzi di trasporto

[E04]Quale mezzo di trasporto utilizzi di più per motivi di lavoro? *

Scegliere la risposta appropriata per ciascun elemento:

	Mai	Raramente	A volte	Spesso	Sempre
Piedi	○	○	○	○	○
Bicicletta privata	○	○	○	○	○
Bike sharing	○	○	○	○	○
Moto	○	○	○	○	○
Tram, pullman, corriera	○	○	○	○	○
Metropolitana	○	○	○	○	○
Car sharing	○	○	○	○	○
Taxi	○	○	○	○	○
Auto privata	○	○	○	○	○
Treno	○	○	○	○	○
Treno ad alta velocità	○	○	○	○	○
Aereo	○	○	○	○	○

[E05]Quale mezzo di trasporto utilizzi di più nel tempo libero? *

Scegliere la risposta appropriata per ciascun elemento:

	Mai	Raramente	A volte	Spesso	Sempre
Piedi	○	○	○	○	○
Bicicletta privata	○	○	○	○	○
Bike sharing	○	○	○	○	○
Moto	○	○	○	○	○
Tram, pullman, corriera	○	○	○	○	○
Metropolitana	○	○	○	○	○
Car sharing	○	○	○	○	○
Taxi	○	○	○	○	○
Auto privata	○	○	○	○	○
Treno	○	○	○	○	○
Treno ad alta velocità	○	○	○	○	○
Aereo	○	○	○	○	○

[E06]Nella tua giornata tipo dell'ultimo mese quale distanza hai percorso dalla tua abitazione a luoghi frequentati per attività di lavoro? *

Scegliere la risposta appropriata per ciascun elemento:

0-1 km	1-5 km	5-20 km	20-50 km	50-100 km	più di 100 km	0-10 minuti	10-20 minuti	20-40 minuti	40-60 minuti	60-90 minuti	più di 90 minuti
○	○	○	○	○	○	○	○	○	○	○	○

Indica sia in temrini di chilometri, sia in termini di tempo impiegato (minuti)

[E07]Ti capita di viaggiare in treno per motivi di lavoro? *

Scegli **solo una** delle seguenti:

- ○ Sì
- ○ No

[E071]Se dovessi andare da casa tua alla stazione, quanti mezzi pubblici potresti prendere? *

Scegli **solo una** delle seguenti:

- ○ nessuno
- ○ 1
- ○ 2
- ○ più di 2

Considerare il numero di alternative possibili.

[E072]Quale mezzo utilizzi di solito per raggiungere la stazione dei treni? *

Scegli **solo una** delle seguenti:

- ○ piedi
- ○ bicicletta
- ○ mezzo pubblico
- ○ motociclo
- ○ auto
- ○ taxi
- ○ Altro

[E073]Quanto dista da casa tua la stazione dei treni? *

Fai una stima dei chilometri

Fai una stima in minuti

Fai riferimento alla distanza casa - stazione.

Se utlizzi meszzi pubblici includi nel conteggio dei minuti: percorso a piedi + attesa del mezzo + percorso con il mezzo

[E074]Utilizzi spesso per spostamenti di lavoro treni ad alta velocità (Freccia Rossa o Italo)? *

Scegli **solo una** delle seguenti:

- ○ Sì
- ○ No

Sezione F. Consumi e Stili di vita

[F01]Sei iscritto/a o partecipi attivamente a *

Scegliere **tutte** le corrispondenti:

- ☐ partito politico
- ☐ sindacato tradizionale
- ☐ associazione quasi sindacale (ACTA, Alta Partecipazione etc.)
- ☐ associazione di categoria professionale
- ☐ movimento civile
- ☐ associazione di volontariato, ONG, ONLUS
- ☐ associazione artistica - culturale
- ☐ nessuna delle precedenti
- ☐ Altro: _____

[F02]In media, al di fuori del tempo dedicato alle attività di lavoro (inclusi gli spostamenti), quanto tempo dedichi alle attività elencate? Indica in ore *

	nei giorni lavorativi	nei giorni festivi
Cura dei figli (o minori conviventi)		
Cura dei genitori		
Lavori di casa		
Cura della propria persona		
Attività formative utili al lavoro		
Attività ricreative esterne al lavoro		

[F03]Nel tempo libero quanto ti piace...? Dai un voto da 1 a 10 *

Scegliere la risposta appropriata per ciascun elemento:

	1	2	3	4	5	6	7	8	9	10
Praticare Sport	○	○	○	○	○	○	○	○	○	○
Dedicarmi ad attività manuali (cucina, bricolage, piccoli lavori di manutenzione, giardinaggio, collezionismo etc.)	○	○	○	○	○	○	○	○	○	○
Frequentare corsi (lingue, strumenti musicali, ballo etc.)	○	○	○	○	○	○	○	○	○	○
Navigare su internet/ usare social network o giochi di ruolo	○	○	○	○	○	○	○	○	○	○
Fare shopping	○	○	○	○	○	○	○	○	○	○
Spettacoli e intrattenimento (teatro, cinema, concerti)	○	○	○	○	○	○	○	○	○	○
Seguire programmi, serie alla tv	○	○	○	○	○	○	○	○	○	○
Fare volontariato/ partecipare a comitati fare politica	○	○	○	○	○	○	○	○	○	○
Leggere	○	○	○	○	○	○	○	○	○	○
Attività all'aria aperta (passeggiate, escursioni, visite dei parchi)	○	○	○	○	○	○	○	○	○	○

[F04]Quando devi prenotare un viaggio per una vacanza... Indica la frequenza dove 1=mai e 10=SEMPRE *

Scegliere la risposta appropriata per ciascun elemento:

	1	2	3	4	5	6	7	8	9	10
Ti rivolgi ad un'agenzia turistica	○	○	○	○	○	○	○	○	○	○
Prenoti autonomamente su internet	○	○	○	○	○	○	○	○	○	○
Prenoti con molto anticipo per trovare tariffe convenienti	○	○	○	○	○	○	○	○	○	○

[F05]Quali aspetti consideri più importanti in una vacanza? Dai un voto da 1 a 10 *

Scegliere la risposta appropriata per ciascun elemento:

	1	2	3	4	5	6	7	8	9	10
Contatto con la natura	○	○	○	○	○	○	○	○	○	○
Strutture attrezzate per famiglie con bambini	○	○	○	○	○	○	○	○	○	○
Vicinanza a casa	○	○	○	○	○	○	○	○	○	○
Impianti sportivi	○	○	○	○	○	○	○	○	○	○
Attrazioni culturali	○	○	○	○	○	○	○	○	○	○
Alloggio di lusso	○	○	○	○	○	○	○	○	○	○
Vita notturna	○	○	○	○	○	○	○	○	○	○
Prezzi bassi	○	○	○	○	○	○	○	○	○	○

[F06]Nei giorni feriali di solito con chi consumi la cena ? Indica la frequenza dove 1=MAI e 10= SEMPRE *

Scegliere la risposta appropriata per ciascun elemento:

	1	2	3	4	5	6	7	8	9	10
da solo/a	○	○	○	○	○	○	○	○	○	○
con familiari	○	○	○	○	○	○	○	○	○	○
con amici/colleghi	○	○	○	○	○	○	○	○	○	○

[F07]Nei giorni feriali di solito come consumi la cena ? Indica la frequenza dove 1= MAI e 10= SEMPRE *

Scegliere la risposta appropriata per ciascun elemento:

	1	2	3	4	5	6	7	8	9	10
Mangio cibo cucinato (da me o da familiari) a casa	○	○	○	○	○	○	○	○	○	○
Compro cibo da asporto	○	○	○	○	○	○	○	○	○	○
Esco per un aperitivo	○	○	○	○	○	○	○	○	○	○
Vado al ristorante / pizzeria	○	○	○	○	○	○	○	○	○	○
Vado a mangiare presso altri (amici / familiari)	○	○	○	○	○	○	○	○	○	○

[F08]In base alle tue necessità, quali sono i servizi più importanti che dovrebbero essere raggiungibili a piedi da casa? Dai un voto da 1 a 10 *

Scegliere la risposta appropriata per ciascun elemento:

	1	2	3	4	5	6	7	8	9	10
asilo nido o scuola materna	○	○	○	○	○	○	○	○	○	○
scuola elementare o scuola media	○	○	○	○	○	○	○	○	○	○
ufficio postale	○	○	○	○	○	○	○	○	○	○
fermata dei mezzi pubblici	○	○	○	○	○	○	○	○	○	○
biblioteca	○	○	○	○	○	○	○	○	○	○
stazione dei treni	○	○	○	○	○	○	○	○	○	○
coworking	○	○	○	○	○	○	○	○	○	○
banca	○	○	○	○	○	○	○	○	○	○
parco / area verde	○	○	○	○	○	○	○	○	○	○
internet point - fotocopisteria	○	○	○	○	○	○	○	○	○	○
supermercato aperto 24/24 ore	○	○	○	○	○	○	○	○	○	○
palestra	○	○	○	○	○	○	○	○	○	○
attività di ristorazione	○	○	○	○	○	○	○	○	○	○
uffici comunali	○	○	○	○	○	○	○	○	○	○

[F09]Indica qual è il livello di accessibilità da casa per ciascuna delle attività elencate *

Scegliere la risposta appropriata per ciascun elemento:

	raggiungibile a piedi	raggiungibile con mezzi pubblici	raggiungibile in macchina	raggiungibile in treno
Fare la spesa (mercati, fiere, supermercati)	○	○	○	○
Mangiare/ bere fuori	○	○	○	○
Andare a vedere uno spettacolo	○	○	○	○
Praticare sport	○	○	○	○
Partecipare a un corso	○	○	○	○
Fare volontariato, fare politica	○	○	○	○

[PREMIO]Per partecipare all'estrazione di una carta regalo Feltrinelli del valore di 100 o 50 euro inserisca un indirizzo mail cui contattarla in caso di vincita *

Scrivere la propria risposta qui:

Il questionario è terminato. **Grazie!**

Potrà partecipare all'estrazione di una carta regalo Feltrinelli del valore di 100 o 50 euro

Per ulteriori informazioni scrivere a claudia.coletta@unimi.it

Inviare il questionario.
Grazie per aver completato il questionario.

Conclusions and discussion

This study is a first attempt to investigate a quite new subject in the labor market field (the Independent PROfessionals, IPROs) by adopting an interdisciplinary perspective with a strong focus on the spatial dimension, often underestimated by sociologists.

Once specified our research interests, we have first designed and carried out a web survey in order to collect first-hand data. Data collection, cleaning and coding process was a long and meticulous one; full details are provided in chapter 2 and 3.

Afterwards we have provided some insights on sample composition by analysing standard socio-economic variables (gender, age, education, income, Isei) and we have tried to test hypothesis related to ICTs use, residential location, work-related commuting habits (chapter 3). Major topics like ICTs frequency use, job-related mobility by frequency and distance have been addressed by computing indexes.
Last part (chapter 4) involved the use of GIS data and it was focused on spatial auto-correlation testing for IPROS residential location.

By analysing our sample of respondents, our major findings in relation to initial research questions are here reported.

Independent professionals (or freelancers) are not an homogenous fast riser class: internal cleavages were detected. Controlling for education level we got evidence of gender (males VS females) and age (youngs 25-35 VS 35-45) segmentation related to annual income levels.
For youngs and women the “status incongruity” theory (Ranci 2012) is a reality: they face income instability despite having high educational profiles (Ranci 2012: 59).

The “Unemplyoyment - push” thesis is supported by the fact that more than 41% of respondents declared their former working status to be *de facto* an unemployment status.

Concerning the spatial dimension: around 89% have experienced residential mobility after 1,5 years of working as Ipros and over 45% of respondents currently live in a city or within its metropolitan area. That is they tend to live in dense central urban nodes where they can easily access advanced services, build social networks and find working opportunities.

Whereas those who do show higher scores on mobility index are more likely to live in smaller municipalities and being females aged between 45-55; they might respond to central areas attracting force, but do not have the chance to live there. What is most interesting is that in such smaller municipalities where Ipros are more mobile satisfaction index measured for local infrastructures

display lower scores. Where it is more needed the infrastructural endowment appear not to be responding to Ipros needs: this could be a first symptom of geographical marginalization phenomena.

As Castells already argued in 1997, the configuration of infrastructure networks is not neutral and contribute to connect or disconnect places. Such aspect should deserve further investigation in order to detect phenomena of spatial exclusion which lead to social ones.

Finally another Castell's thesis (1996) is confirmed: mobility index was proved to be positively correlated with ICTs use index. In other words ICTs use "does not imply the end of the office, but the diversification of working sites for a large fraction of the population, and particularly for its most dynamic, professional segment".

Last point involved detecting clusters of Ipros on the territory.

Due to many missing data for analysis made at municipality level and including Southern areas (few or no respondents were located in such areas), we decided to performed the analysis at province level from North to Centre Italy. Thus data were collected at municipality level, but aggregation was made at province scale.

The 'Modifiable Areal Unit Problem' (MAUP) has been recognized by geographic literature as one of the biggest arbitrary decision influencing findings; this was a big limit of our research: it made us loose details and implied choosing bigger territorial unit of analysis.

This strong limit influenced our findings and almost no evidence was found of clustering patterns according spatial auto-correlation tests performed on different neighbourhood structures and adjacency matrixes. The polycentric city region model could not be retrieved in the North west of Italy, neither in other geographical areas and few so-called "central areas" (main urban nodes of the region) were detected by making reference to main macro economic variables mentioned in chapter 4: Torino, Genova, Milano, Bologna, Firenze, Roma.

Nevertheless the relevance of some macro economic variables in working as a pulling factors for Ipros in so called central areas was confirmed by GLM analysis: rate of medium-high specialised jobs; number of Active local companies; number of research centres, number of patent applications, average income per capita.

To conclude is extremely important to underline the weaknesses and the strenghts of this study.

The lack of official data sources available on the object under investigation and the lack of resources to perform survey sampling mainly influenced the whole research design and findings.

Main overall limitations of the research design involved the absence of a filtering sampling method and the reliance on the web channel as the main collector of answers. The self-selected sample therefore is not supposed to be a representative subset of Ipros population and any findings we have described are specific for the respondents' sample and no statistical inference can be made.

The strength of this dissertation relies on providing relevant theoretical contributions to economic sociology and labour studies by highlighting the existence of new subjects such as the professional freelancers and proposing an alternative perspective for studying sociological phenomena from a spatial point of view.

Mobility is the key dimension as it could be regarded at as a spatial as well as social phenomenon.

Throughout the dissertation we tried to highlight the role of the space in influencing living and mobility trends and eventually drawing patterns of segregation and exclusion.

Finally the aim of this research is mainly an explorative one on a new sociological subject (quantifying and describing a phenomenon). The hope is that this dissertation heralds the beginnings of a wider research agenda into this important but neglected aspect of the independent professional workforce.