VOLUME LXIII – NN. 1-2 GENNAIO- GIUGNO 2009

RIVISTA ITALIANA DI ECONOMIA DEMOGRAFIA E STATISTICA

Direzione, Redazione e Amministrazione
Piazza Tommaso de Cristoforis, 6
00159 ROMA
TEL. e FAX 06-43589008
E-mail: sieds@tin.it

Stampato da CLEUP sc
"Coop. Libraria Editrice Università di Padova"
Via G. Belzoni, 118/3 – Padova (Tel. 049/650261)
www.cleup.it

INDICE

IN THIS ISSUE

This volume collects the invited papers of the XLVI Scientific Meeting of the Italian Society of Economics, Demography and Statistics (SIEDS) held in Florence on May 28-30, 2009. The Meeting was organized together with the Department of Statistics "Giuseppe Parenti", University of Florence, and had "Poverty and Social Exclusion" as main topic.

The Meeting was opened by Professor Giovanni Maria Giorgi who presented a brief historical survey on the first 70 years of the SIEDS.

The main lecture was delivered by the Honorary President of the Society, Professor Luigi Di Comite. He focussed on "Migration and Social Marginalization", highlighting some aspects concerning the territorial mobility of populations over the last one hundred and fifty years from a temporal viewpoint as well as across the EU countries – particularly in its old configuration of 15 member states – by a territorial perspective. Due to its specific feature of touching on several aspects of the Meeting's main topic, this paper has been included at the beginning of this volume.

Moving on a brief description of the various subjects dealt with during the Meeting, it seems important to note that all the other papers have been published following the alphabetical order of the first Author and not the order of presentation in which they were delivered.

The paper by Gianni Betti, Achille Lemmi and Vijay Verna provides an overview of the methodologies and applications used in the context of the EU for the definition of regional indicators of poverty and social exclusion.

Claudio Ceccarelli and Giovanni Maria Giorgi used the so-called Analysis of Gini (ANOGI) on data from Italian survey on income and living condition (It-Silc) for investigating whether attrition introduces an element of bias in the analysis of Italian income distribution at the regional level.

Enrica Chiappero Martinetti compares the approaches of Amartya Sen and John Roemer, with reference to the concept of poverty as lack of possibility and lack of opportunity respectively, thus pointing out analogies and differences between these points of view.

Starting from the Gini coefficient, Conchita D'Ambrosio's paper introduces other indices to be used for the evaluation of deprivation and social exclusion, also highlighting their links with social well-being.

On the basis of the outcomes of a survey carried out by "Banca d'Italia" and using two different approaches for the standardization of income levels, Guido Ferrari and Mauro Maltagliati compare the households welfare according to territorial divisions as well as socio-economic profiles.

The paper by Enrico Giovannini, Adolfo Morrone and Giulia Ranuzzi provides the guidelines of the OECD Project known as “Measuring the Progress of Societies” and aims at evaluating whether any consensus on how welfare and progress can be measured is actually emerging.

The paper delivered by Vincente Gozàlvez Pérez and José Ramón Valero Escandrell focuses on social inequalities with regard to the risk of poverty in the regions of Spain, particularly referring to single-person households and the Gypsies.

Highlighting some of the outcomes of the application of a model, the paper by Massimiliano Mascherini and Andrea Saltelli aims at providing an identikit of the active citizen in Europe.

The study presented by Daria Mendola, Annalisa Busetta and Anna Maria Milito proposes a new index – validated and tested on a sample of individuals participating in an ECHP (European Community Household Panel) research project – which measures poverty in a longitudinal perspective.

Fausta Ongaro and Silvana Salvini analyse the relationship between health status and socio-economic conditions in the light of the most recent data on health conditions of the Italian population. The paper aims at evaluating whether the differences in health levels observed in Northern and Southern Italy may even depend on socio-economic factors.

Monica Pratesi, Caterina Giusti and Nicola Salvati resort to M-quantile small area estimation methods to estimate the Head Count Ration, the mean income and some income quantiles in the Provinces of three Italian Regions. They also estimate the cumulative distribution function of the income in the principal Municipalities of these Regions.

Finally, the paper delivered by Giambattista Salinari and Gustavo De Santis deals with the main features of the so-called “modified Gibrat’s model” – i.e. a method useful to describe the evolution of income, with particular attention to small incomes.

Enrico Del Colle

RIEDS Editor

INTRODUCTION TO THE XLVI SCIENTIFIC MEETING OF THE ITALIAN SOCIETY OF ECONOMY, DEMOGRAPHY AND STATISTICS

Giovanni Maria Giorgi
SIEDS President

I do not want to hide here my satisfaction to hold the XLVI Scientific Meeting of the Italian Society of Economics, Demography and Statistics at the prestigious Florentine Athenaeum where, 70 years ago, our Society saw the light due to the wish and the hard work of Livio Livi and other eminent scholars.

In the last 70 years, our Society operated to actively contribute advances in several fields, such as statistics, demography, social and economic disciplines, in order to fulfil the statutory goals.

The official birth of the Society dates back to a decree signed by the King of Italy, Vittorio Emanuele III, in 29th June 1939 and published in the Official Gazette of 11th September 1939, by which the Advisory Committee for the Study of Population was transformed into the Italian Society of Demography and Statistics (SIDS), located in Florence, Viale Curtatone n.1.

Initially, the Society published the proceedings of scientific meetings and only later was endowed with its own journal called Italian Review of Demography and Statistics, whose first issue was published in January 1947.

Later, at the meeting of Society members of 18th April 1950, the request of prof. Luigi Amoroso and other scholars to extend the Society scientific interest to economics was granted. Our association took its current name of "Italian Society of Economics, Demography and Statistics" and, in the same year, the journal was renamed "Italian Review of Economics, Demography and Statistics" (RIEDS).

Scientific reasons were definitely prevalent in giving life to an Italian Society; however, taking into account the strong nationalism of that period, someone remarked that "it was also a matter of national pride".

As a matter of fact, although in Italy there was an established tradition for demographic and statistical studies, there was not a scientific association that allowed a close and organic link between scholars who met, in most cases, abroad, in scientific meetings organized by the International Statistical Institute and the International Union for the Study of Population.

At that time, as Livio Livi emphasized in the introduction to the first meeting of our Society (held at ISTAT in Rome on May 27-28, 1939), there were already 16 Societies of Statistics in Europe, as well as similar Societies were operating in Canada, USA, Mexico, Brazil, Argentina, British India, Japan and China.

This explains why there was full agreement among scholars, when the Advisory Committee for Population Studies was turned into a scientific society.

Nevertheless, "*creating a Society of Demography only would have been as to do something incomplete*" as Livio Livi stressed. A Society ensuring the possibility of an active collaboration to all the scientific tendencies was needed in order "*to do a wide national-based organism and not a group bound to a given place or a given school*".

The adhesions to SIDS were many and, among the first ones those of Rodolfo Benini and many other distinguished academics.

In fact, on the occasion of the scientific meeting held in Florence (December 28, 1940) the SIDS had 122 ordinary members, among them Carlo Emilio Bonferroni whose contributions in statistics and mathematics are still of importance (e.g., his index of concentration and his famous "inequalities"), Raffaele D'Addario who gave important methodological contributions both to the study of the distribution of income and wealth and to the solution of particular actuarial problems, Mario de Vergottini, and above all we cannot forget Giuseppe Parenti who was also the Rector of the University of Florence, Giuseppe De Meo (future Istat President), Pierpaolo Luzzatto Fegiz, Diego De Castro, Guglielmo Tagliacarne, Bruno Grazia Resi, Ugo Papi (future Rector of "La Sapienza" University of Rome) and Silvio Vianelli.

These scholars played a prominent role in the diffusion of Statistics, Demography, Economics and Social Statistics and gave important scientific contributions, some of which had exceeded the national borders and even today, as mentioned above, are still highly topical.

We have briefly recalled our origins since without an aware memory we would be without future.

This meeting wants to be both a celebration but also a stimulus to face problems of a complex present.

Despite the great economic difficulties faced by scientific institutions in our country, we look forward with hope to continue the revitalization of SIEDS at both national and international level; the presence in this scientific meeting of prestigious Italian and foreign speakers (among others, Anthony Atkinson) is a clear mark of this fact.

The theme chosen for our meeting, "*Poverty and social exclusion*", has to be considered as an attempt to reflect together on how to tackle, from a scientific point of view, issues of great social and economic relief.

Therefore, consistently with the spirit of the Society founders and, above all, with the auspice for a future of better perspectives for the scientific research of our country, we declare open the XLVI Scientific Meeting of the Italian Society of Economics, Demography and Statistics.

Florence, 28th May 2009

Rivista Italiana di Economia Demografia e Statistica Volume LXIII nn. 1-2 – Gennaio-Giugno 2009

MIGRATION AND SOCIAL MARGINALISATION*

Luigi Di Comite

1. Introduction

During the centuries, peoples have regularly moved over the land urged by the need to improve their living conditions (Simon, 2008), but the migratory phenomenon – which should be analyzed both from the flows and stocks point of view at the same time – acquired a particular relevance more or less a century ago, whether, above all in the past, migrants came from (comparative) densely populated areas and went towards (comparative) less populated areas, or they have lately left from developing countries or from countries which are at an in-between condition of development, to go towards fully developed countries.

With reference to these last decades experience, it can be stated that a typical background of the migratory phenomena is the one which involves the generally well-pondered decision to migrate abroad often alone at the beginning, with a chosen destination – both inside and abroad – to settle down, probably using contacts already in situ. From this point of view, in many cases, they are completely or partially joined by their families to find themselves in the immigration country usually in a situation of social and economic marginalization, which often partly clashes with the one they had in the origin areas.

Comparing to these ones (and to their part), the immigrants have better economic incomes, which go along with a social marginalization, that can sometimes become "ghettoization".

Shortly, they leave a situation in which they are fully or at least enough integrated in the social structure (even if in not satisfying economic conditions) to settle in a new territory in a condition which generally: a) often causes only a partial integration in the receiving area from a social point of view; b) has a dualistic aspect from an economic point of view because: i) there is an income which is higher – often remarkably higher – than that before the migration, which

* Work made within the framework of the research program "Caratteristiche evolutive dei processi demo–economici e mobilità territoriale delle popolazioni. Gli attuali equilibri in alcune macro–aree regionali", financed by Università di Bari (year 2009) and directed by prof. Luigi Di Comite; translated by dr. Luisa D'Aniello.

allows also to give money to the family in the origin country; ii) the income is generally rather inferior compared to that of the native population, also when the immigrants do jobs which are not desirable and that are dangerous and heavy.

That being stated, not having now the real possibility to make an in-depth analysis of this last aspect of the phenomenon, in the following pages particular attention will be directed only to international migration, leaving out the internal ones for the moment.

2. Population territorial mobility

During the centuries – and also the millenniums – the human race has always moved over the territory for a variety of reasons, that can be more or less traced back to what we can consider the need of a "better living condition".

Since the second half of the XIX century, this attitude to mobility [1] has interested a great number of migrants and has started to set up short-range migrations – which took the form of internal migrations – against long-range migrations, which were often transoceanic and interested first of all Europe as origin areas and then America (United States, Brasil, Argentina) as destination countries.

Nowadays, migratory flows – including return migrations – have a global aspect and every country (of suitable demographic dimensions) is interested both in outgoing and incoming flows.

In terms of stock, on the other hand, at the beginning of the third millennium (see Table 1), about 190,5 millions of people – more or less the 3,0% of the global population – lived in a country different from that of origin[2] (United Nations, 2009), more (115 millions) in the developed countries and less (75 millions) in the developing countries[3]. No country in the world escaped this phenomenon, not even the least developed countries even if for these ones, as shown in the above mentioned Table 1, the foreign presence was less remarkable in 2005 than in 1990, in opposing trend.

---

[1] The idea we are now referring to is that of not-forced mobility: i.e. not determined by push-factors, such as environmental disasters, slave trade, etc..

[2] "Foreign" country should mean a country other from that (or those) of which a person has the right to have the passport.

[3] A rather small number of foreigners (or more precisely of "non-native") staying in the developing countries lived (see Table 1) in the countries which are considered the least developed in the world (most of all African countries).

Table 1 – *Foreign population (millions of migrants) by continents and major areas, 1990-2005.*

Continents	Foreign population		Increment	Percentage	
	1990	*2005*	*1990-2005*	*1990*	*2005*
Africa	16,4	17,1	0,7	10,6	9,0
Asia	49,8	53,3	3,5	32,2	28,0
Latin America and Caribbean	7,0	6,6	-0,3	4,5	3,5
U.S.A. and Canada	27,6	44,5	16,9	17,8	23,3
Europe	49,4	64,1	14,7	31,9	33,6
Oceania	4,8	5,0	0,3	3,1	2,6
World	**154,8**	**190,6**	**35,8**	**100,0**	**100,0**
More developed countries[4]	82,4	115,4	33,0	53,2	60,5
Less developed countries	72,5	75,2	2,8	46,8	39,5
Least developed countries[5]	*11,0*	*10,5*	-0,5	*7,1*	*5,5*

Source: United Nations (2009)

[4]They are: 1) European countries; 2) North America (United States and Canada); 3) Japan; 4) Australia and New Zealand.

[5]The group of least developed countries currently comprises 50 countries: Afghanistan, Angola, Bangladesh, Benin, Bhutan, Burkina Faso, Burundi, Cambodia, Cape Verde, Central African Republic, Chad, Comoros, Democratic Republic of the Congo, Democratic Republic of Timor-Leste, Djibouti, Equatorial Guinea, Eritrea, Ethiopia, Gambia, Guinea, Guinea-Bissau, Haiti, Kiribati, Lao People's Democratic Republic, Lesotho, Liberia, Madagascar, Malawi, Maldives, Mali, Mauritania, Mozambique, Myanmar, Nepal, Niger, Rwanda, Samoa, Sao Tome and Principe, Senegal, Sierra Leone, Solomon Islands, Somalia, Sudan, Togo, Tuvalu, Uganda, United Republic of Tanzania, Vanuatu, Yemen and Zambia.

Moreover, the above mentioned data arouse some perplexities about the "role" played by the stowaways – permanent and/or transitory – that are generally remarkable in number and "hidden", both in the developed countries and in the near developing ones and that should be completely excluded from the evaluations when such data are used.

In many countries of the world, the foreign population (born abroad) had a rate higher than 10,0% and these countries – including different ones of rather small demographic dimensions – were larger in number in 2005 than in 1990 (see Figure 1).

In truth, until 2005 only 8 countries with a population with more than 20 million inhabitants had a foreign presence higher than the rate of 10,0%[6] and there were also some other countries of modest demographic dimensions where the foreign presence was higher than 30,0%[7].

The ten countries with the highest rate of foreign presence[8] included more than the half of its total amount. Furthermore, this rate was higher in 2005 (53,6%) than in 1990 (52,1%), attesting the attitude towards the territorial concentration of this phenomenon. Moreover, in 2005 there were also the United Kingdom and Spain among those countries, confirming the role of the European Union in the migratory framework. The UK and Spain had replaced Pakistan and Canada.

[6]They are: 1) Saudi Arabia (25,9%), 2) Australia (20,3%), 3) Canada (18,9%), 4) Ukraine (14,7%), 5) U.S.A. (12,9%), 6) Germany (12,3%), 7) Spain (11,1%) e 8) France (10,7%): in 2005, 83,2 million foreigners (i.e. the 43,7% of the total amount of the "foreign population") lived globally in these eight countries.

[7] Among these countries there are, for example, Qatar (78,3%), United Arab Emirates (71,4%), Kuwait (62,1%), Luxemburg (37,3%), Bahrein (35,8%).

[8] Eight of these countries (U.S.A., Russian Federation, Germany, Ukraine, France, Saudi Arabia, Canada, India and UK) have not changed between 1990 and 2005: only Pakistan and Canada were replaced by UK and Spain in 2005.

Table 2 – *The 30 countries with the highest number of foreign population, 1990-2005.*

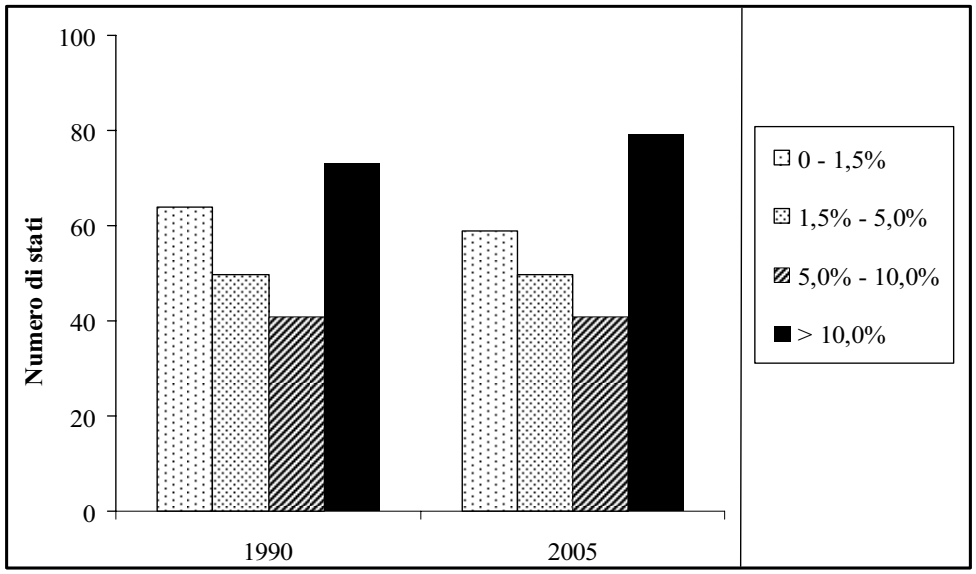
Rank	1990			2005		
	Countries	Foreign population		Countries	Foreign population	
		Number of migrants (millions)	As percentage of total		Number of migrants (millions)	As percentage of total
1	U.S.A.	23,3	15,0	U.S.A.	38,4	20,2
2	Russian Federation	11,5	7,4	Russian Federation	12,1	6,4
3	India	7,4	4,8	Germany	10,1	5,3
4	Ukraine	7,1	4,6	Ukraine	6,8	3,6
5	Pakistan	6,6	4,2	France	6,5	3,4
6	Germany	5,9	3,8	Saudi Arabia	6,4	3,3
7	France	5,9	3,8	Canada	6,1	3,2
8	Saudi Arabia	4,7	3,1	India	5,7	3,0
9	Canada	4,3	2,8	United Kingdom	5,4	2,8
10	Australia	4,0	2,6	Spain	4,8	2,5
11	Iran	3,8	2,5	Australia	4,1	2,2
12	United Kingdom	3,8	2,4	Pakistan	3,3	1,7
13	Kazakhstan	3,6	2,3	United Arab Emirates	3,2	1,7
14	China	2,2	1,4	China	3,0	1,6
15	Ivory Coast	2,0	1,3	Israel	2,7	1,4
16	Uzbekistan	1,7	1,1	Italy	2,5	1,3
17	Argentina	1,6	1,1	Kazakhstan	2,5	1,3
18	Israel	1,6	1,1	Ivory Coast	2,4	1,2
19	Kuwait	1,6	1,0	Jordan	2,2	1,2
20	Switzerland	1,4	0,9	Japan	2,0	1,1
21	Italy	1,3	0,8	Iran	2,0	1,0
22	United Arab Emirates	1,3	0,8	Singapore	1,8	0,9
23	Sudan	1,3	0,8	Palestine	1,7	0,9
24	Belarus	1,3	0,8	Ghana	1,7	0,9
25	South Africa	1,2	0,8	Kuwait	1,7	0,9
26	Netherlands	1,2	0,8	Switzerland	1,7	0,9
27	Malawi	1,2	0,8	Malaysia	1,6	0,8
28	Ethiopia	1,2	0,8	Netherlands	1,6	0,8
29	Turkey	1,2	0,8	Argentina	1,5	0,8
30	Jordan	1,1	0,7	Turkey	1,4	0,7
-	***Total***	***116,3***	***75,1***	***Total***	***146,9***	***77,1***
-	**World**	**154,8**	**100,0**	**World**	**190,6**	**100,0**

Source: see Table 1

As to the situation of the whole word, the data in Table 2 show not only the globalization of the phenomenon, but also some relevant aspects such as:

a) the clear predominance, as immigration areas, of two particular territorial areas, i.e. North America (U.S.A and Canada) and the European Union – above all if considered in its old composition of 15 countries[9] - , with six countries (Germany, France, United Kingdom, Spain, Italy and Netherlands) which are present in Table 2, also with Luxembourg, where the "not-native" were the 37,3% of the entire population in 2005;

Figure 1 – *Distribution of countries by the percentage of migrants in the population, 1990 and 2005.*



b) the presence of the two most important States of U.R.S.S., i.e. the Russian Federation and Ukraine, even if this last one is better known as emigration rather than immigration country;

c) the marginal position of Australia and above all South Africa;

[9] The 15 States on the European Union were: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, Sweden and United Kingdom.

d) the almost total absence of countries from Latin America and Caribe, since in the above mentioned chart there is only Argentina: this is limited to 2005 and in a position of evident rearguard;
e) the characteristics of the Chinese and Indian situations, since in these two countries – both with a population which greatly exceeds one billion people – the remarkable amount of foreign population (in decline in India and growing in China) is due above all to their huge demographic dimension;
f) the African case, where the predominant country is from the francophone west part of Africa (Ivory Cost), which has long attracted the population from near countries.

3. Migrations between the end of the XIX century and the beginning of the XXI century

As intended nowadays, it can be stated that the great migratory flows started not before the second half of the XIX century and concerned above all European countries as origin countries and the New World (America and Oceania) as destination countries (see Figure 2).

They were essentially consistent wide-range population movements that – involving some dozen million people – mainly interested relatively over-populated countries as origin areas and as destination countries the ones where the autochthonous population was slowly marginalized and sometimes drastically cut down in number because of the clash with the new arrived population

From the beginning of the XIX century to the clash of the First World War, the demographic trend of the different European countries was characterized by a constant widening of the difference between birth and death rates[10].

The resulting demographic explosion, in its turn, caused – also for reasons of balance – an answer in terms of population territorial mobility, in the sense that migrations, generally the transoceanic ones, helped containing (or also avoiding) the effects of the constant growth of the rate of natural increase of the European countries population.

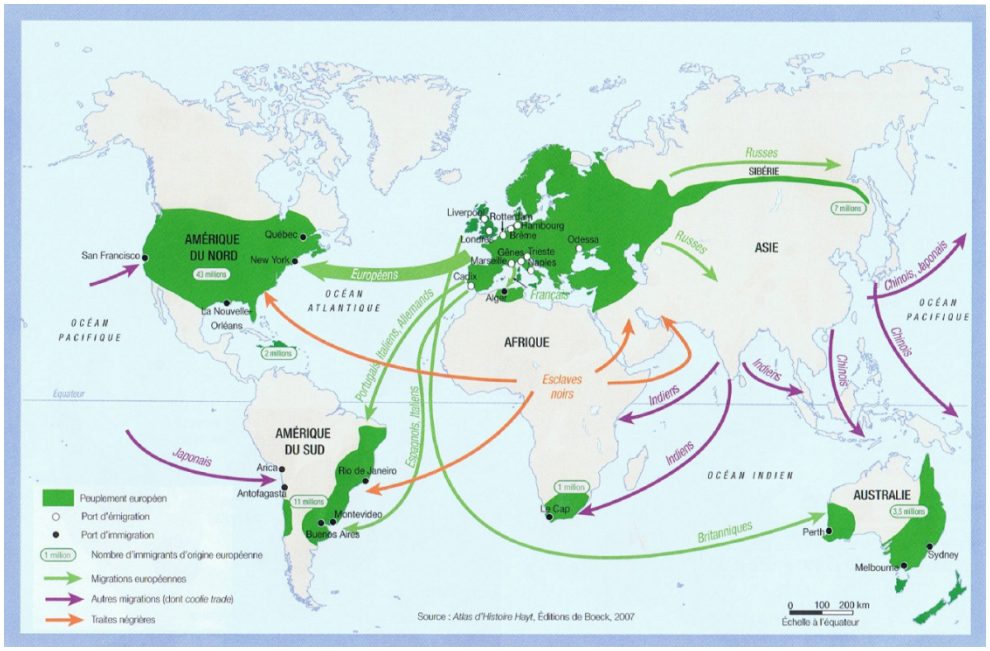
[10] Essentially, it was the beginning – and the continuation – of the processes of (first) demographic transition (Di Comite, 1980; Chesnais, 1986), that have characterized the European demography for more than a century: more or less, in the different territorial areas, from the first half of the XIX century to the second half of the XX century.

As it can be seen in Figure 2, almost up to the beginning of the First World War the favorite destination for the migratory flows coming from the European countries were:

a) the United States of America with the Southern part of Canada; b) a part of Latin America and precisely the cost part of Brazil, Uruguay and Argentina, together with the Northern part of Chile; c) South Africa; d) Oceania (Australia and New Zeeland). At the beginning, the migrants came above all from countries of Northern Europe (United Kingdom, Germany, Scandinavian countries, etc…); later on they came from countries of the Mediterranean Europe – which were less developed from a social-economic point of view (above all Italy and Spain) –.

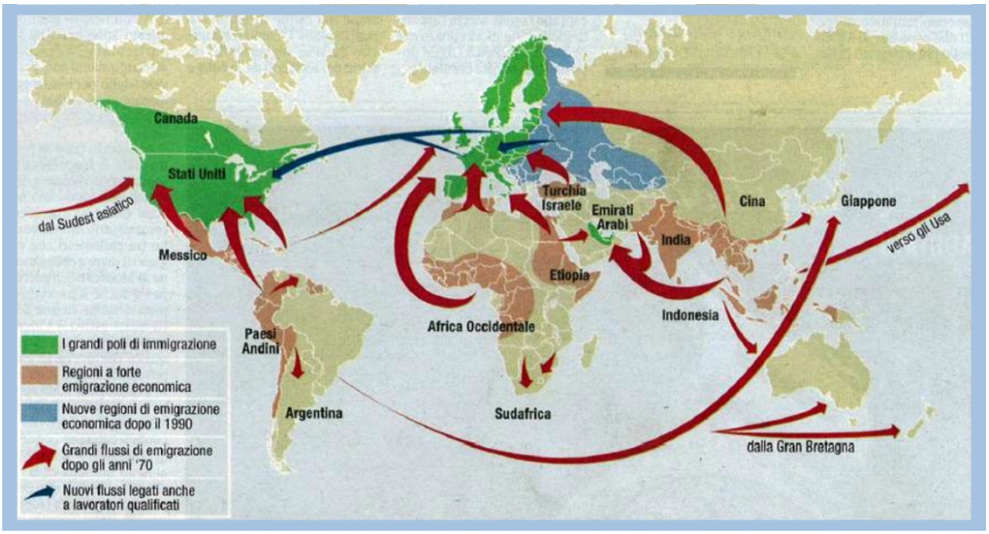
Nowadays, or better between the last decades of the XX century and the first decade of the XXI century, the framework of reference of population territorial mobility has considerably changed, above all for international migrations, and the recent migratory flows are involving rather all the world, as origin, transition and destination areas (see Figure 3).

Figure 2 – *Main international migratory flows of the XIX century.*



Source: G. Simon (2008)

Figure 3 – *Main international migratory flows between the end of the XX and the beginning of the XXI century.*



In fact, the international migratory "marketplace" has widened and specialized. This justifies the existence and relevance of particular figures of migrants, such as the so called political and environmental refugees. However, the most part of the actual migratory flows is based mainly on economic reasons[11] and, generally, it is oriented along a "path" which has the developing countries as origin areas and the developed countries as final destination. The developing countries are still characterized by both a transitional demography and a slow social-economical development, while the developed ones are mainly characterized by a demographic stagnation as well as a high social-economical development[12].

Today, one of the main areas of interest for the present migrations – both legal and illegal – is the European Union: within its 27 member states, foreign people (non-native) are quite the 8,0% of the population, while within the 15 states configuration the rate is about 10,0%.

[11] This is true for the so called "family reunion".
[12] Among these ones, there are the countries which are members of the European Union and the ones of the 15-State European Union.

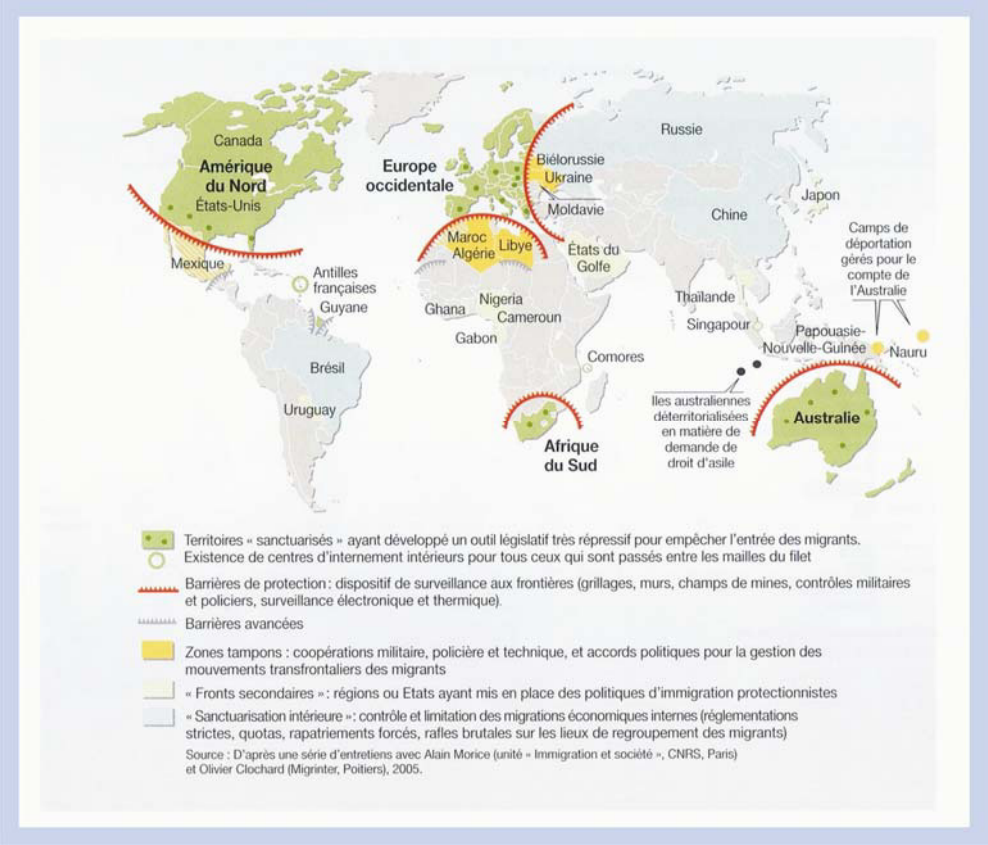
4. Migrations concerning European Union

As to the territorial mobility of population between the end of the XX century and the beginning of the XXI century, the two main attracting areas for the international migratory flows (see Figure 4) were North America (U.S.A. and Canada) and the European Union, above all if considered in its 15-State configuration. Leaving out North America, we have to pay attention to the European Union and to what has happened in this area for the last thirty years. At the end of the 70s, its southern part (or Mediterranean) started to turn from a traditional emigration area to an area of recent immigration.

At the end of the 80s, more or less together with the fall of the Berlin wall, within the 15-State European Union there were States with a strong migratory tradition such as Benelux, France, Germany and the United Kingdom, all situated in the Northern part of the EU together with States of recent or even very recent immigratory experience, all within the Mediterranean area of the EU (Italy, Spain, Portugal and Greece). The former ones had a presence of foreigners higher than the rate of 5,0%; in the latter ones the presence did not reach the rate of 2,0%. Moreover, some of these Mediterranean countries – precisely Italy and Spain – have been for a long time the main "entrance doors" for migrants who, coming from both the southern part of the Mediterranean basin (above all from Mediterranean and non Mediterranean Africa), and near Balkans as well as far east countries, wanted to be accepted in the different countries of the EU, often living submerged.

At the beginning of the 90s, the EU countries of greatest migratory impact were France, Germany and the United Kingdom, thanks to their large demographic dimension. The ethnical features of the foreigners living in these countries were very different (see Figure 5): in Germany there was a predominance of Turkish people, followed by a small amount of Yugoslav people; in France there was a predominance of Maghribi – together with them we can consider also those born in Maghreb become naturalized French (Pellicani – D'Addato, 2004) – followed at a distance by Portuguese people. In the United Kingdom there was a strong predominance of extra-European nationalities.

Figure 4 – *Contact areas of the main international migratory flows between the end of the XX century and the beginning of the XXI century.*



Source: see Figure 2

There are evident differences within the above mentioned states, in their common situation of great migratory countries. There are even greater differences if we compare the situation of these countries with that of the Mediterranean ones of the EU on one hand and with the situations of the remaining countries of the 15-State EU and of the other twelve on the other, with reference to the beginning of the 90s.

Nowadays, i.e. after less than twenty years, the situations have deeply changed, both on the global and regional level (for example, the EU, the Mediterranean Basin, Latin America, etc.)

First of all, it can be easily observed how the phenomenon has widened: for example, from data in Table 2 it is clear that on a global level, the foreigners (not native) have passed from 154,8 million people in 1990 to 190,6 million people in 2005, with a growth of 23,1% of the initial amount[13].

As to both the EU and the Mediterranean Basin, form the data of the above mentioned Table, importance must be given to Spain that with 4,8 million foreigners in 2005 – with a foreign incidence of 12,0% – is the tenth country in the list of the immigration countries while it had a foreigner presence inferior to 2,0%[14] fifteen years before.

Within the Mediterranean countries of the EU, Spain and Italy are now among the main receivers of the traditional migratory flows, along the South-North and East-West lines; moreover, Spain has been the main receiving country of the flows from Latin America for many years.

Spain and Italy have the common problem of the illegal arrivals, which generally take place within the area of the Gibraltar Straits and/or the West African lane (Morocco and Mauritania) – Canary Islands for Spain, Sicily Channel but also Otranto Channel for Italy – with relevant cases of human deaths (Simon, 2008).

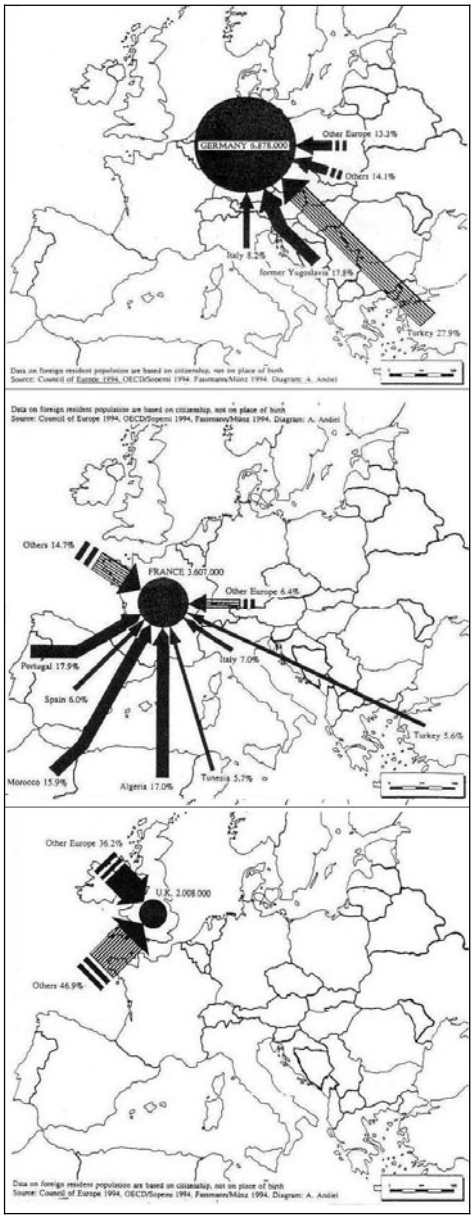
For both these countries, it is important to stress that the foreigner presence contributes to the actual growth of fertility, as the rate of children born from mothers who are foreigner citizens has increased in the last years compared to that of foreigners, stowaways included (Di Comite – Girone, 2007).

The last consideration is the possibility for the EU to have a common immigration policy, also possible for the old 15-State configuration and even more possible for the 27-State configuration.

[13] At the same time, it must not be forgotten that, during these 15 years, the world population would have passed from 5.290,5 million inhabitants in 1990 to 6.512,3 million inhabitants in 2005, with an increase of 23,1%, exactly alike the observed one for the migratory stock.

[14] In fact, since 1990 until today, Italy (above all the South), Spain, Portugal and Greece have slowly changed from emigration into immigration areas.

Figure 5 – *Nationality of foreign people living in Germany, France and United Kingdom at the beginning of the 90s.*



Source: Münz (1995)

After observing Figure 5, it is important to consider the other countries of the EU: just considering the different situations from a country to another gives the possibility to have a negative answer because needs and expectations deeply vary from countries of ancient migration to countries of actual migrations, ending with the ones that could become countries of immigration in the future.

5. Conclusions

A lot of pages have been written about migratory phenomena, strictly intended in their different demographic meanings and about their consequences in the demographic and economic field. There are different points of view and different time and territorial references in this variety of works.

We have tried to make a summary – surely not exhaustive – of the topic analysed, with the aim of reaching an old target never got before, i.e. the chance to study migrant people, comparing their previous situations in the origin countries with the ones (often temporary) in the receiving or settling countries, considering that we should speak about marginalization or even social exclusion rather than acceptance of migrant people.

Leaving out also the recent past, we have to pay attention to what has happened between the end of the XX and the beginning of the XXI century, or more exactly between 1990 – after the fall of the Berlin wall – and nowadays.

Between 1990 and 2005 (see Table 1) the amount of foreign people[15] living all over the world increased of about 35 million units, with an increase of 23,1% compared to the first data and at the same time (see reference 13) world population underwent a similar increase: it should be said that there is a substantial balance between the two events (increase in the number of world population and increase of "not native" population). Really, it is not so: the increase in the number of foreign people is limited to two territorial areas (North-America and EU), which have both high economic development and demographic stagnation due to low – often very low – fertility rates. In these countries, immigration has a role of replacement (Pellicani, 2002), compared with birth deficit, but it sometimes causes refusal phenomena, which affect above all the weakest members, first of all gipsy people and the so-called "sans papiers".

[15] It is important for the reader to pay attention again about the fact that in this Table, foreign population regards the "non-native" population, i.e. people born abroad.

References

ACOCELLA N. – SONNINO E. (a cura di) (2003), *Movimenti di persone e movimenti di capitale in Europa,* il Mulino, Bologna.
AMBROSINI M., MOLINA S. (a cura di) (2004), *Seconde generazioni. Un'introduzione al futuro dell'immigrazione in Italia,* Fondazione Agnelli, Torino.
BARBAGLI M. (1998), *Immigrazione e criminalità in Italia,* il Mulino, Bologna.
BARSOTTI O. (1998), *Migrants: a Resource for the Country of Destination and a Development Factor for the Country of Origin,* "Atti della XLII Riunione Scientifica della Società Italiana di Statistica", Sessioni plenarie- Sessioni specializzate, Bari, 9-11 giugno.
BIANCO M. L. (1996), *Classi e reti sociali. Risorse e strategie degli attori nella riproduzione delle disuguaglianze,* il Mulino, Bologna.
BIANCO M. L. (a cura di) (2001), *L'Italia delle disuguaglianze,* Carocci, Roma.
BONIFAZI C. (1998), *L'immigrazione straniera in Italia,* il Mulino, Bologna.
BONIFAZI C., CONTI C., GESANO G. (2000), *Dinamica demografica e migrazioni nell'area del Mediterraneo,* in GOMEL G. e ROCCAS M. (a cura di), *Le economie del Mediterraneo,* Banca d'Italia, Roma.
CAGIANO DE AZEVEDO R. (2000), *Le migrazioni internazionali. Il cammino di un dibattito,* Giappichelli, Torino.
CASTLES S., KOSACK G. (1976), *Immigrazione e struttura di classe in Europa Occidentale,* F. Angeli, Milano.
CHESNAIS J.C., (1985), *Progrès économique et transition démographique dans les pays pauvres: trente ans d'expérience (1950-1980),* "Population", n° 1.
CHESNAIS J.C., (1986), *La transition démographique. Etapes, formes, implications éecomiques,,* "Travaux et Documents", Cahiers n° 113, Presses Universitaires de France, Paris.
M.C., CONIGLIO N., FERRI G. (2007), *L'esercito degli invisibili. Aspetti economici dell'immigrazione clandestina,* il Mulino, Bologna.
COLASANTO M., AMBROSINI M. (a cura di) (1993), *L'integrazione invisibile. L'immigrazione in Italia tra cittadinanza economica e marginalità sociale,* F. Angeli, Milano.
COLEMAN D. (2006), *Immigration and Ethnic Change in Low-Fertility Countries: A Third Demographic Transition?,* "Population and Development Review", n° 3.
COLOMBO A., SCIORTINO G. (a cura di) (2002), *Assimilati ed esclusi,* il Mulino, Bologna; C ORTI P. (2003), *Storia delle migrazioni internazionali,* Laterza, Bari.
DAL LAGO A. (1999), *Non-persone. L'esclusione dei migranti in una società globale,* Feltrinelli, Milano.
DE SANTIS G. (1997), *Demografia ed economia,* "Studi e ricerche", il Mulino, Bologna.
DI COMITE L. (1980), *Teoria e prassi della transizione demografica,* Studi in onore di Paolo Fortunati, vol. I, CLEUB, Bologna.
DI COMITE L. (1981), *Problemi statistici delle migrazioni,*, "Annali di Statistica", Serie IX, vol. I, n. 4.

DI COMITE L. (1983), *L'emigrazione italiana nella prima fase del processo transizionale,* "Giornale degli economisti", n° 7-8.
DI COMITE L. (1985), *Aspetti dell'emigrazione italiana negli Stati Uniti (1880-1897),* Conference on "Societies in Transitions: Italians and Italian-Americans in the 1980's", Balch Institute-Temple University, Philadelphia, 11-12 october 1985, in "Analisi storica", n. 5.
DI COMITE L. (1986), *Aspects of Italian Emigration, 1881-1915,* in I. A. GLAZIER & L. DE ROSA (eds.), *Migration across Time and Nations,* Holmes & Neier, New York.
DI COMITE L. (1993), *Transizione demografica e fenomeni migratori,* "Poligonos", Universidad de Leon, n° 3.
DI COMITE L. (1995) *La présence maghrébine dans les pays de l'Europe méditerranéenne,* in L. DI COMITE (a cura), *Le migrazioni maghrebine,* Quaderno n° 9 del Dipartimento per lo Studio delle Società Mediterranee dell'Università di Bari, Bari.
DI COMITE L., BONERBA P., GIRONE S. (2006), *La popolazione. Tra stagnazione e sviluppo*, in MALANIMA P. (a cura di), *Rapporto sulle economie del Mediterraneo – Edizione 2006,* il Mulino, Bologna.
DI COMITE L., GIRONE S. (2007), *El papel de la inmigración extranjera en la recuperación de la fecundidad en algunos países mediterráneos de la Unión europea,*"Nike", n° 2-3.
DI COMITE L. - GLAZIER I.A. (1984), *Socio-Demographic Characteristics of Italian Emigration to the United States from Ships Passengers Lists: 1880-1914,* "Ethnic Forum", n° 1-2.
DI COMITE L. - MORETTI E. (1996), *Demografia e flussi migratori nel Bacino mediterraneo,* 2.a ristampa, La Nuova Italia Scientifica, Roma.
DI COMITE L., PATERNO A. /a cura/ (2001), *Quelli di fuori. Dall'emigrazione all'immigrazione: il caso italiano*, F. Angeli, Milano, 2001.
FAKIOLAS R., KING R. (1996), *Emigration, return, immigration: a review and evaluation of Greece's post-war experience of international migration*, International Journal of Population Geography.
FARGUES PH., (1992), *Interpréter la démographie: quel rôle tient-elle dans la peur du lendemain?*, in DI COMITE L. – IAQUINTA P. (a cura di), *Demografia e demo-economia del Bacino mediterraneo*, Quaderni del Dipartimento per lo Studio delle Società Mediterranee dell'Università di Bari, n° 5, Cacucci, Bari.
FARGUES PH. - PELLICANI M.C. (2000), *La possibile influenza del contesto demografico sulla mobilità del fattore umano: un'analisi per generazioni,* "Rivista Italiana di Economia, Demografia e Statistica", n° 3.
FELD S. et MANÇO A. (2000), *L'intégration des jeunes d'origine étrangère dans une société en mutation,* L'Harmattan, Paris.
GOLINI A. (2004), *Globalization, Demographic Trends, and Human Mobility,* "Atti della XLII Riunione Scientifica della Società Italiana di Statistica", Sessioni plenarie- Sessioni specializzate, Bari, 9-11 giugno.
GOLINI A. /a cura/ (2006), *L'immigrazione straniera: indicatori e misure di integrazione,* il Mulino, Bologna.

HAUG W. - COURBAGE Y. - COMPTON P. (1998), *Les caractéristiques démographiques des minorités nationales dans certains Etats européens,* Etudes démographiques, n° 30, Conseil de l'Europe, Strasbourg.
KING R. (1986), *Return Migration and Regional Economic Problems*, Coom Helm, Kent.
ISHIKAWA Y and MONTANARI /edited by/ (2003), *The New Geography of Human Mobility. Inequality Trends?,* Società Geografica Italiana, Rome.
LIVI BACCI M. (2002), *Storia minima della popolazione del mondo,* Il Mulino, Bologna.
MACDONALD J.S. (1964), *Chain migration, ethnic neighbourhood formation and social networks,* "The Milbank Memorial Fund Quarterly", vol. XLII, n° 1.
MACDONALD J.S. (1992), *Chain Migration Reconsidered,* "Bollettino di Demografia Storica", n° 16.
M. MACURA–COLEMAN D. /eds./ (1994), *International Migration: Regional Processes and Responses,* United Nations Economic Commission for Europe-United Nations Population Fund, "Economic Studies", n° 7, United Nations, New.
MA MUNG E. (2000), *La diaspora chinoise: géographie d'une migration,* Editions Ophrys, Paris.
MINGIONE E. (1999), *Le sfide dell'esclusione: metodi, luoghi, soggetti,* il Mulino, Bologna;
MONALDI V. (1996), *I paesi meno avanzati e i programmi d'aggiustamento strutturale: un'analisi preliminare,* in G. ANCONA – F. BOTTA (a cura), *Problemi di cooperazione e sviluppo nel Bacino mediterraneo,* Quaderno n° 12 del Dipartimento per lo Studio delle Società Mediterranee dell'Università di Bari, Bari.
MORLICCHIO E. (2000), *Povertà ed esclusione sociale. La prospettiva del mercato del lavoro,* Edizioni Lavoro, Roma.
MÜNZ R. (1995), *Where did They All Come from? Tipology and Geography of European Mass Migration in the Twentieth Century*, Evolution or Revolution in European Population, European Population Conference, Milano 4-8 sept. 1995, F. Angeli, Milano.
NEGRI N., SARACENO C. (2000), "Povertà, disoccupazione ed esclusione sociale", "Stato e mercato", n. 2, pp. 175-210.
PACI M. /a cura di/ (1993), *Le dimensioni della disuguaglianza,* il Mulino, Bologna.
PELLICANI M.C. (2002), *Transizione demografica, invecchiamento della popolazione e migrazioni di sostituzione,* in L. DI COMITE L. – A. PATERNO (a cura), *Quelli di fuori. Dall'emigrazione all'immigrazione: il caso italiano*, F. Angeli, Milano, 2001.
PELLICANI M.C., D'ADDATO A.V. (2004), *Modalità di ottenimento della cittadinanza francese ed analisi quantitativa delle principali caratteristiche dei francesi di acquisizione,* in G. ANCONA – R. PACE (a cura), *Mobilità, partecipazione e sviluppo,* Quaderno n° 31 del Dipartimento per lo Studio delle Società Mediterranee dell'Università di Bari, Bari.
POLLINI G., SCIDÀ G. (2004), *Sociologia delle migrazioni e della società multietnica,* F.Angeli, Milano.
PUGLIESE E. (2002), *L'Italia tra migrazioni internazionali e migrazioni interne,* Universale Paperbacks, il Mulino, Bologna.
SALVINI S. (2004), *Demographic Convergence and Economic Divergence? Revisiting Demographic Transition in Developing Countries,* "Atti della XLII Riunione Scientifica

della Società Italiana di Statistica", Sessioni plenarie- Sessioni specializzate, Bari, 9-11 giugno.
SANCHEZ BARRICARTE J.J. (2008), *El crecimiento de la población mundial :implicaciones socioeconómicas, ecológicas y éticas,* Tirant lo blanch, Valencia.
SARDON J.P. (2004), *Evolution démographique récente des pays développés,* "Population", n° 2.
SIMON G. (2008), *La planète migratoire dans la mondialisation,* A. Colin, Paris.
STROZZA S. (2008), *Partecipazione e ritardo scolastico dei ragazzi stranieri e d'origine straniera,* "Studi emigrazione", n° 171.
THUMERELLE P. J. (1996), *Les populations du monde,* Éditions Nathan, Paris.
UNITED NATIONS (2009), *International Migration Report 2006: a Global Assessment,* New York.
VERON J.(1994), *Population et développement,* Presses Universitaires de France, Paris.
VERON J.(2004), *Sustainable Development: a North-South Issue?,* "Atti della XLII Riunione Scientifica della Società Italiana di Statistica", Sessioni plenarie - Sessioni specializzate, Bari, 9-11 giugno.
ZANDATO D., CASSATA L., MARTIRE F., STROZZA S., VITIELLO M. (2008), *L'integrazione come processo multi-dimensionale. Condizioni di vita e di lavoro degli immigrati,* "Studi emigrazione", n° 171.
ZANFRINI L. (2004), *Sociologia delle migrazioni,* Laterza, Bari.
ZANFRINI L. (2004), *Sociologia della convivenza interetnica,* Laterza, Roma-Bari.
ZINCONE L. (1994), *Uno schermo contro il razzismo,* Donzelli, Milano.

SUMMARY

In these few pages are reported some short considerations about the territorial mobility of populations mainly concerning the past one hundred and fifty years according to temporal terms, as well, the European Union countries – particularly in its old configuration of 15 members – by a territorial perspective, as together to the North American countries of (U.S.A. and Canada), they both are still the favourite destinations of the recent migration flows.

Luigi DI COMITE, Dipartimento per lo Studio delle Società Mediterranee (DSSM) – Università di Bari.

Rivista Italiana di Economia Demografia e Statistica Volume LXIII nn. 1-2 – Gennaio-Giugno 2009

MULTIDIMENSIONAL ASPECTS OF POVERTY AND LIVING CONDITIONS: THE ESTIMATION AT REGIONAL LEVEL

Gianni Betti, Achille Lemmi, Vijay Verma

1. Introduction

Indicators of poverty and social exclusion are an essential tool for monitoring progress in the reduction of these problems. Such indicators are usually calculated at national level, since their construction is based on sample surveys that rarely have a statistical significance at regional level. In the EU-wide context, these indicators need to be comparable across countries and time. For this purpose, the European Commission has adopted a common set of indicators, referred to as the Laeken Indicators. A critical review of these indicators provides the starting point of our research (Section 2). Hitherto, most of these indicators have been defined only at the national level. These are not necessarily appropriate or sufficient for regional analysis. In Section 3, we make specific proposals for going from the country list to a regional list of indicators.

The strategy we have adopted for the construction of regional indicators has three fundamental aspects (Section 4): making the best use of available survey data; exploiting to the maximum 'meso' data; and using the two sources in combination to produce the best possible estimates for regions using appropriate small area estimation (SAE) techniques. Implementation of SAE methodology is described in Section 5. The main application involves the production of estimates of country-to-NUTS1 and NUTS1-to-NUTS2 ratios. A very small selection of the results on indicators at NUTS2 level is shown in Section 6.

2. Laeken indicators of poverty and social exclusion

Indicators of poverty and social exclusion are an essential tool for monitoring progress in the reduction of these problems. In the EU-wide context, these indicators are most useful when they are comparable across countries, so that the situation in any EU Member State can be evaluated in relation to the situation in other countries. These indicators also need to be comparable over time for monitoring trends. In order to meet these objectives, the European Commission has adopted a common set of indicators of poverty and social exclusion, with

standardised definitions and procedures for their construction. These are referred to as the Laeken Indicators, deriving their name from the location of the European Council meeting where they were adopted. The set of common indicators is supplemented by country-specific indicators, chosen flexibly according to the requirements and data availability in individual countries. Hitherto, most of the indicators have been defined and constructed only at the national level, except for occasional breakdown for special subpopulations such as children, other groups by age and gender, or different household types.

Our main interest in this paper is to define indicators at the regional (sub-national) level, and describe and illustrate the statistical methodology for their construction.

Some of the country-level indicators can be usefully classified down to the regional level in their existing form; some other may need modification (simplification) before such classification. There are, however, also country-level indicators which are not suitable (meaningful, useful, feasible) for regional breakdown. It is also necessary to consider additional, specifically regional indicators which are not covered in the country-level list. Still, the established set of country-level indicators provides the basis for developing indicators suitable for the regional level. Therefore, we begin by describing in this section the indicators defined for country-level application. Selection, adaptation, and supplementation of these indicators for regional application will be taken up in the next section.

3. Adaptation to the regional level

The first issue in developing regional indicators concerns the choice of the type of units to serve as 'regions'. For a number of substantive and practical reasons, we consider geographical-administrative regions, specifically NUTS regions (and LAUs) at various level of classification, as the most appropriate choice for EU countries. The average population size for EU countries is somewhat over 5 million for NUTS1, of the order of 1.8 million for NUTS2, and 375 thousand for NUTS3 regions. The units vary considerably in size across the countries. However, generally the range of variation across countries declines as we go down the hierarchy. When measures at the regional level are constructed by aggregating information on individual elementary units, two types of measures which can be so constructed should be distinguished:

- Average measures, i.e. ordinary measures such as totals, means, rates and proportions constructed by aggregating or averaging individual values. (Examples: area unemployment rate; population proportion having a certain characteristic).

o Distributional measures, such as measures of variation or dispersion among households and persons in the region; such measures depend on the distribution of the whole population.

The patterns of variation and relationship for the two types of measures can differ from each other, and hence require separate statistical models. Average measures are often more easily constructed or are available from alternative sources. Distributional measures tend to be more complex and are less readily available from sources other than complex surveys; at the same time, such measures are more pertinent to the analysis of poverty and social exclusion.

An important point to note is that, more than at the national level, many measures of averages can also serve as indicators of disparity and deprivation when seen in the regional context: the dispersion of regional means is of direct relevance in the identification of geographical disparity.

3.1 *Cross-sectional measures of income poverty*

Basic indicators. Henceforth the Laeken indicators have been applied primarily at the national level. It is necessary to adapt them for regional application, taking into account differences in the requirements and the data situations. As a general rule, it is necessary to focus on the more basic among the indicators. This is because the data requirements are substantially increased when the results are to be geographically disaggregated. Detailed disaggregation of the indicators by age, gender and other characteristics - simultaneously with disaggregation by geographical region – has to be severely restricted, especially when the information comes from sample surveys of limited size (say less than 2,000 sample households per region). Broad classification, such as distinguishing children, youth and elderly persons, may be possible, but even that has to be subsidiary to the need for adequate regional breakdown. For the purpose of regional indicators, the focus has to be primarily on ordinary poverty rates for the total population, possibly with some major breakdowns. Certain more complex poverty and inequality measures - measures which are more sensitive to details and irregularities of the empirical income distribution - are less suited to disaggregation to small populations and small samples. Examples are Gini coefficient, relative median at-risk-of-poverty gap, and at-risk-of-poverty rate before social transfers.

On the other hand, poverty rates have to be supplemented by other indicators not considered explicitly in the Laeken list. Perhaps the most important of these is simply the mean income levels of the regions, the dispersion among which provides a measure of regional disparities. General entropy measures may also be

useful because they can be decomposed into within and between region components.

Poverty line levels. By the "level of poverty line" we mean the population level to which the income distribution is pooled for the purpose of defining the poverty line. All poverty related indicators in the Laeken list are based on country poverty lines (for instance as 60% of the national median income). This applies even when the indicators are aggregated over countries, or are disaggregated to regions within a country. Usually, the income distribution is considered separately at the level of each country, in relation to which a poverty line is defined and the number (and proportion) of poor computed. These numbers may then be pooled over countries to obtain the EU poverty rate (but still defined in terms of national poverty lines). Similarly, we may disaggregate the numbers poor by region and obtain regional poverty rates defined according to the national poverty line in each country.

It is also useful to consider poverty lines at other levels. For instance, we may pool the data across countries to construct a single income distribution (and hence a single poverty line) for the whole EU, and use this to compute poverty rates at the EU level, or for individual countries, or for any level of regions within any country. Specifically relevant for constructing regional indicators is the use of regional poverty lines, i.e. a poverty line defined for each region based only on the income distribution within that region. The numbers of poor persons identified with these lines can then be used to estimate regional poverty rates. They can also be aggregated upwards to give alternative national poverty rates, or disaggregated downwards to produce sub-regional poverty rates – but in all cases they are based on the regional poverty lines. The poverty line level chosen can make a major difference to the resulting poverty rates when it is higher than the level of analysis or aggregation. The extent depends on the degree of disparity between the units of analysis. However, we find that the poverty line level chosen often makes only a small difference to the resulting poverty rates when it is the same as or lower than the level of analysis or aggregation. For instance, while country poverty rate for a country can differ greatly depending on whether a EU poverty line is used, the country rates tend to differ much less whether we use a poverty line defined at the national, NUTS1 or NUTS2 level.

3.2 *Indicators of non-monetary deprivation*

In addition to the level of monetary income, the standard of living of households and persons can be described by a host of indicators, such as housing conditions, possession of durable goods, the general financial situation, perception of hardship, expectations, norms and values. The data required for the construction

of non-monetary indicators are generally simpler to collect than detailed data on monetary incomes. This makes such indicators more convenient and suitable for regional analysis. An index of non-monetary deprivation which summarises a range of indicators of living conditions should be developed and analysed in its own right. Using the methodology described in Betti and Verma (2008), we have constructed measures of overall non-monetary deprivation.

It is also useful to combine monetary and non-monetary measures in order to study the extent to which they overlap. If individuals are subject both to income poverty and non-monetary deprivation simultaneously, their overall deprivation is more intense. Similarly, if they are subject to only one of the two, their deprivation is, in relative terms, less intense.

We have also constructed an alternative version of monetary poverty using the fuzzy set approach (for a review of the overall approach see Lemmi and Betti, 2006). In this approach, poverty is seen as a matter of degree determined by the individual's location in the income distribution. For a description of the methodology, see Giorgi and Verma (2002).

4. Methodology

The strategy recommended in this research for the construction of regional indicators of poverty and deprivation has three fundamental aspects:

- making the best use of available sample survey data, such as by cumulating and consolidating the data to construct more robust measures which can permit a greater degree of spatial disaggregation;
- exploiting to the maximum 'meso' data (such as highly disaggregated tabulations available in NewCronos) for the purpose of constructing indicators for small areas;
- using the two sources in combination to produce the best and most complete possible estimates for sub-national regions using appropriate small area estimation (SAE) techniques.

4.1 *Survey data*

To explain the statistical procedures in concrete terms, it is useful to consider the actual data sets we have utilised. The main results are based on European Community Household Panel (ECHP). For EU-15 countries, we have 8 annual waves of comparable, longitudinal microdata (income reference years 1993-2000), except for some initial waves missing in three countries. Single rounds of similar,

but less comparable, microdata are available for two other countries (Poland and Romania). For a number of other (mostly new EU) countries, we have no microdata, but only published aggregate indicators on poverty. The microdata contain information on equivalised household disposable income and sample weights, so that diverse measures of poverty and inequality can be estimated. The same applies to non-monetary indicators of deprivation, though here the data are less complete. One critical shortcoming of the data is that regional identifiers are not available at all, or are available only for NUTS1 level in many cases.

Where the information comes from sample surveys of limited size, a trade-off is required between temporal detail and geographical breakdown. Generally, the different ECHP waves provide a consistent and comparable series and the results can be averaged over waves to increase precision, that is, to increase the effective sample size. Of course, the core of the sample is a panel of the same individuals so that data from the different waves are highly correlated. Therefore the effective sample size for estimates averaged over waves is increased by much less than proportionally to the number of waves included. Nevertheless, there is a significant increase in the effective sample size due to real variation over time in the composition of the sample, in characteristics of the individuals and households, and also due to the presence of response variability and other random effects.

In the standard analysis, as for instance in Laeken indicators, the poverty line is defined as a certain percentage (x%) of the median income of the national population. By "poverty line threshold" we mean the choice of different values of 'x'. The Laeken set includes a measure of dispersion around the at-risk-of-poverty threshold (computing the percentage of persons, over the total population, with an equivalised disposable income below, respectively, 40%, 50%, 60% and 70% of the national median equivalised disposable income). The substantive objective of introducing indicators of dispersion around the poverty line is to take more fully into account differences among countries in the shape at the lower end of the income distribution. Lower thresholds isolate the more severely poor and tend to be more sensitive in distinguishing among countries or other population groups being compared. As the threshold is raised, this sensitivity generally tends to fall: clearly in the extreme case when 'x' is taken as 100% (poverty line equal to the median), the poverty rate in all situations is 50%, by definition.

In addition to the above systematic differences, the results from using different poverty line thresholds are also likely to be affected by irregularities in the empirical income distribution. Irregularities are larger when the distributions are estimated from smaller samples, as normally is the case for disaggregated estimates by region. It is this consideration which is likely to dominate in the context of constructing regional measures. Some gain in sampling precision can be obtained by computing poverty rates using different thresholds, and then taking their

weighted average using some appropriate pre-specified (i.e., constant or external) weights. This is the strategy we have used in the construction of regional indicators, in preference to constructing separate indicators for different thresholds.

4.2 *NewCronos*

The NewCronos (now termed "Eurostat Free Dissemination Database") provides a valuable data resource for the construction of regional indicators. In itself it is not a source of original data, but a compilation of information from a diversity of sources presented in the form of very detailed tabulations. NewCronos REGIO domain covers the principal aspects of the economic and social life of the European Union: demography, economic accounts, labour force, health, education, etc., by region. The concepts and definitions used are as close as possible to those used by Eurostat for the production or compilation of statistics at national level. The standard model for compiling regional aggregates at various levels has been as follows: first, data from various national sources are compiled in the National Statistical Offices, and then provided to Eurostat for validation. This data set is then loaded into NewCronos by the thematic unit in question (in our case the REGIO domain of NewCronos). Most of indicators are available at NUTS2 level.

There are three main forms in which we have utilised variables derived from NewCronos for the construction of regional indicators.

Some statistics in NewCronos can serve, in their own right, as direct indicators pertaining to poverty and living conditions. In fact, the scope for such use is likely to be greater in the context of regional indicators, compared to that in the national context. This is because measures of levels – which are more abundantly available in NewCronos than the generally more complex distributional measures - can themselves serve as indicators of disparity when compared across regions.

A large number of measures correlated with direct indicators of poverty and deprivation can be constructed. In conjunction with direct indicators obtained from more intensive surveys, these measures can be used as "covariates" or "regressors" to produce more precise indicators using small area estimation (SAE) procedures described in Sections 4.3 and 5.

In addition, NewCronos provides a very large number of measures, giving what has been termed as "intermediate output" indicators. Such indicators express on the one hand the policy effort in favour of those at risk of poverty and social exclusion, and on the other hand the impact of social policies as well as of the economic context. NewCronos is a unique source of such indicators.

We believe that this resource, NewCronos, has hitherto been under-utilised, and that there is a great potential for more thorough exploitation of the information

which already exists. While direct indicators of regional poverty and living conditions are generally not available with sufficient regional breakdown in NewCronos, several exceptionally positive aspects of the resource need to be appreciated. Some of these become even more important as we move down from the national to the regional level.

- o A wide range of subject-matter areas are covered in the very detailed tabulations provided. These can be utilised to construct many direct indicators pertaining to poverty and living conditions, as well as to obtain many more variables correlated with direct indicators.
- o Detailed break-down – especially for variables correlated with direct indicators of interest – is available, mostly to NUTS2, and in a few cases to NUTS3 level.
- o NewCronos is a dynamic resource, in principle regularly updated as new or improved data become available. Of course, its timeliness, statistical quality and comparability depend on the original data sources from which the information is derived. But the very process of bringing those data into a unified framework through a centralised operation can be expected to enhance data quality in all its dimensions – coherence, consistency, completeness, transparency, and also comparability.
- o The data base is accessible and convenient to use, and most importantly, this resource is placed in the public domain as Eurostat Free Dissemination Database.

4.3 *Small area estimation (SAE)*

In the paper we take the view that rather than discussing the SAE procedures in general terms independent of the actual data situation, it is more useful to develop and implement the estimation procedures in concrete terms on the basis of the data sources presently available to us. Such a practical approach is much more likely to bring out the variety of situations and problems actually encountered in the course of producing regional indicators of poverty and deprivation.

In the literature small area models are classified as: (i) area level random-effect models (Fay and Herriot, 1979), which are used when auxiliary information is available only at area level (such as the prevailing unemployment rate); (ii) nested error unit-level regression models, used if unit specific covariates (such as the individual's or the household's employment situation) are available at unit level (Battese *et al.*, 1988).

In our main application we apply area level random-effect models relating small area direct estimates to domain specific covariates, considering the random area

effects as independent. The basic area-level model includes random area specific effects, and the area specific covariates, $x_i=(x_{i,1}, x_{i,2}, \ldots, x_{i,p})$, are related to the target parameters θ_i (totals, means, proportion, etc.) as follows $\theta_i = x_i\beta + z_i\nu_i$ with i = 1…m, where z_i are known positive constants, β is the regression parameters vector px1, ν_i are independent and identically distributed random variables with 0 mean and variance σ_ν^2. Moreover it is assumed that the direct estimators $\hat{\theta}_i$ are available and design unbiased, in the form $\hat{\theta}_i = \theta_i + e_i$, where e_i are independent sampling errors with zero mean and known variance ψ_i. The Best Linear Unbiased Predictor (BLUP) estimator of θ_i is:

$\tilde{\theta}_i\left(\sigma_\nu^2\right) = x_i\hat{\beta} + b_i^T GZ^T V^{-1}\left(\hat{\theta}_i - x_i\hat{\beta}\right)$, where bit is a mx1 vector (0,0,…,0,1,0,..,0) with 1 referred to the i-th area and β are estimated by generalized least square as $\hat{\beta} = \left(X^T V^{-1} X\right)^{-1} X^T V^{-1}\hat{\theta}$.

The BLUP estimator is a weighted average of the design-based estimator and the regression synthetic estimator: $\tilde{\theta}_i\left(\sigma_\nu^2\right) = \gamma_i\hat{\theta}_i + \left(1-\gamma_i\right)x_i\hat{\beta}$, where: $\gamma_i = \dfrac{\sigma_\nu^2}{\sigma_\nu^2 + \psi_i}$ is a weight (or 'shrinkage factor') which assumes values in the range [0-1]. This parameter measures the uncertainty in modelling θ_i. (Gosh and Rao, 1994). Mathematical details for the BLUP estimators are available in Handerson (1950). An important point to note is that the mean square error of the BLUP estimator depends on the variance parameter σ_ν^2, which in practice is unknown and is replaced by its estimator $\hat{\sigma}_\nu^2$, so that a two stage estimator $\tilde{\theta}\left(\hat{\sigma}_\nu^2\right)$ is obtained; it is called Empirical BLUP (EBLUP).

In Section 4.1 we have described the data situation for the present research in terms of availability of micro-level survey data. In view of this data situation, the options we have considered are summarised in Table 1. Two distinct approaches are involved, depending on the data situation.

(1) EBLUB models

With a few thousand households observed in a sample survey, most estimates at the national level are sufficiently accurate (have small sampling error) to be directly reported. Below the national level, we have used area-level EBLUB composite estimators in countries where available data permitted that, that is, where area-coded survey data are available. In these countries, the following applies.

At NUTS1 level, the available sample sizes are generally smaller and consequently sampling errors are larger than at the country level. In some cases, the NUTS1 samples are very small, and significant gains in precision are obtained by using composite estimates. However, overall the introduction of modelling and composite estimation adds only marginally to the precision of the direct estimates from the survey at NUTS1 level, especially when data can be cumulated over time to enhance the available sample size, as in the present case.

The gains from modelling are naturally more significant at NUTS2 level, and substantially more so at NUTS3 level. Note that NUTS3 are not always 'small' areas in terms of population size; their smallness in the SAE methodology refers to the smallness of the samples available for direct estimation.

Table 1 – *Structure of the modelling.*

	Data situation	Type of estimator used
(1)	Access to area-coded survey data + Access to area-level covariates +Unclustered samples[#]	Composite (area-level EBLUP)
(2)	Lack of access to area-coded survey data, or access only to country-level survey estimates + Access to area-level covariates	Synthetic (regression-prediction)

[#] "Unclustered sample" simply means a sample where the primary sampling units (PSUs) are confined to be within (or at least to coincide with) the areas for which estimates are required. This is generally the case in EU surveys even for NUTS3 regions.

5. SAE model implementation

Each of the two situations identified in Table 1 types involves estimating several models: for each dependent variable of interest, one model corresponding to a particular level of regions, such as 'Model 1' corresponding to EBLUP model for NUTS1 regions, 'Model 2' corresponding to NUTS2 regions, etc. The important point to note is that in estimating each model, information is pooled over a set of countries. This implies the assumption of similar relationships between the model variables in different countries. This is a strong assumption. Its justification results from the 'ratio approach' we have adopted in the model specification, as explained below. The SAE approach we have adopted may be considered somewhat simplistic in that it does not attempt to incorporate temporal or spatial autocorrelations. A major positive feature of the approach, however, is that the

modelling strategy is designed to be hierarchical. We begin with poverty rates and other target variables at the national level, using essentially direct survey estimates without involving any modelling.

We can expect the predictive power of the model at the regional level to be substantially improved when the target variables as well as the covariates are expressed in terms of their values at the preceding higher level. Thus for NUTS1 region i, all target variables and all covariates in the model are expressed in the form of the ratio $R_i = Y_i / Y_0$, where (Y_i, Y_0) refer to the actual values of the variables, respectively, for NUTS1 i and its country. In this way the effect of the difficult-to-qualify institutional and historical factors, common to the country and its regions, is abstracted. This makes the pooling of data across different countries for the estimation of a common model more reasonable. Similarly, in going from NUTS1 region i to its NUTS2 region j, we express the model variables in the form $R_{ij} = Y_{ij} / Y_i$; and similarly from NUTS2 to NUTS3 in the form $R_{ijk} = Y_{ijk} / Y_{ij}$. This type of modelling is further improved by taking different parts of a large or exceptionally heterogeneous country as separate units, examples being eastern and western parts of Germany, or the northern and southern parts of Italy. The same may apply to metropolitan versus other areas in some countries, such as the UK and France. The same ideas are extended to the modelling of subpopulations, such as children, old persons, single person households, etc. We simply model the ratio of the subpopulation measure to the total population measure.

As noted, two different types of SAE models have been estimated:

SAE Model 1: estimated on the ratio NUTS1/Country;

SAE Model 2: estimated on the ratio NUTS2/ NUTS.

One such model has been estimated for each target variable at each level; all countries with area-coded survey data and the particular target variable available are pooled together for the estimation of model parameters at the level concerned.

In this and the next Sections, we describe the models and results for Models 1 and 2. Corresponding to Models 1 and 2, simple regression-prediction models have been used in countries or regions where no area-coded survey data are available. One such model corresponds to each SAE model; it uses the regression coefficients determined from the corresponding SAE model (for one set of countries), to predict the target variables (for another set of countries) on the basis of available predictors. Table 2 lists the 13 target variables for Models 1 and 2. The variables were described in Sections 3.1 and 3.2 and are grouped into three sets: income poverty related measures; overall deprivation measures; and dimension-specific deprivation measures.

Table 2 – *Target variables for SAE models 1 and 2.*

Income poverty related measures		
1	HCR_C	Head Count Ratio – country poverty line
2	HCR_N2	Head Count Ratio – NUTS2 poverty line
3	LogIncPC	Mean log(per capita income)
4	LogEqInc	Mean log(equivalised income)
5	FM_C	Fuzzy monetary poverty rate
Overall deprivation measures		
6	FS_C	Fuzzy supplementary (non-monetary) deprivation rate
7	LAT_C	Latent deprivation rate
8	MAN_C	Manifest deprivation rate
Dimension-specific deprivation measures		
9	FSUP-1	Deprivation rate: dimension 1 (basic life-style);
10	FSUP-2	Deprivation rate: dimension 2 (secondary life-style);
11	FSUP-3	Deprivation rate: dimension 3 (housing facilities);
12	FSUP-4	Deprivation rate: dimension 4 (housing deterioration);
13	FSUP-5	Deprivation rate: dimension 5 (environmental problems);

As far as measures FM_C and FS_C are concerned, the availability of these variables in the countries is as follows. Mostly, the variables are available for EU-15 countries. Sufficient information is not available in the ECHP surveys in Germany, Luxembourg and Sweden to construct deprivation measures in specific dimensions (variables F SUP 1-5). Only monetary measures could be computed from the survey in Romania. It should also be noted that some of the non-monetary measures for Poland lack comparability with corresponding ECHP measures because of differences in the survey questions used.

For countries other than EU15, Poland and Romania, we have no micro data available and only two of the target variables could be constructed from published data (in Eurostat publications Statistics in Focus and also recorded in NewCronos): head count ratio with country poverty line, and log equivalised income.

With regards to the choice of the independent variables (covariates) for building the models, they were selected if the required data were available for most of the countries involved in estimation and regression models. Substantive considerations were also involved in the selection of the covariates used. We decided to estimate the models considering the full set of covariates available and selected, including some which were statistically non-significant.

5.1 *SAE Model 1*

This concerns EBLUP models for going from country to NUTS1 level, used in combination with survey data and the information compiled in NewCronos.

As explained above, we used the 'ratio approach' to improve the precision of the models. In this approach, the model input consists of

(a) NUTS1-to-Country ratio for the statistic concerned, as directly estimated from the survey

(b) standard error of this ratio estimator

The output from the model consists of

(a)* model estimate of NUTS1-to-Country ratio for the statistic concerned

(b)* mean-squared error of this estimate

Performance measures

Table 3 shows some performance measures of SAE Model 1. For each model (i.e., each target variable), three measures are shown:

- the model parameter gamma (γ). It is the ratio between the model variance and the total variance, and is the share of the weight given to the direct survey estimate in the final composite estimate;
- ratio (a)*/(a), i.e., the ratio between the EBLUP estimated value (a)* and the corresponding direct estimate (a). This is to check the extent to which the modelling changes the input direct estimates;
- ratio (b)*/(b), i.e., the ratio between mean-squared error (MSE) of the EBLUP estimate of the NUTS1: Country ratio, and MSE of direct survey estimate of this ratio. This is to check the extent to which the modelling has improved precision of the estimates.

For each of the above, the following summary statistics are given: the mean value over all NUTS1 areas in the model; the coefficient of variation of those values; and the minimum and maximum values. The regression coefficients and the associated significance levels were also estimated in order to evaluate the performance of the models. These may be found in Verma *et al.* (2005).

Overall, the results are as expected: the SAE Model 1 for NUTS1 level does not provide much gain, as can be seen from the mean ratio of mean-squared errors. This is because the sample sizes for most NUTS1 areas are actually quite large; NUTS1 can hardly be called 'small areas' in most ECHP surveys. Further increase in the effective sample size was achieved by cumulation of data over survey waves.

Table 3 – *Performance measures for SAE Model 1. (gamma value, ratio of EBLUP estimates to direct estimates, ratio of EBLUP standard error to direct standard error).*

	Gamma				Estimate EBLUP/direct estimate				Mean-squared error (MSE) MSE(EBLUP)/MSE(direct estimate			
	mean	CV	min	max	mean	CV	min	max	mean	CV	min	max
1 HCR_C	0,86	0,15	0,41	0,99	0,99	0,10	0,70	1,49	0,95	0,19	0,35	1,90
2 HCR_N2	0,35	0,47	0,03	0,73	1,00	0,05	0,84	1,14	0,67	0,23	0,23	0,93
3 logEqInc	0,95	0,05	0,71	0,99	1,00	0,00	1,00	1,00	0,98	0,02	0,89	1,00
4 logIncPC	0,95	0,05	0,71	0,99	1,00	0,00	1,00	1,00	0,98	0,02	0,89	1,00
5 FM_C	0,83	0,16	0,35	0,98	0,99	0,05	0,72	1,05	0,92	0,07	0,68	0,99
6 FS_C	0,83	0,16	0,39	0,98	1,00	0,05	0,84	1,28	0,93	0,07	0,70	0,99
7 Latent	0,86	0,14	0,38	0,98	1,00	0,03	0,81	1,11	0,94	0,06	0,70	0,99
8 Manifest	0,66	0,36	0,15	0,96	0,98	0,12	0,60	1,39	0,83	0,18	0,43	0,99
9 Fsup_1	0,93	0,05	0,74	0,99	1,00	0,02	0,96	1,03	0,97	0,02	0,89	1,00
10 Fsup_2	0,86	0,10	0,65	0,98	1,00	0,03	0,89	1,11	0,94	0,04	0,84	0,99
11 Fsup_3	0,70	0,32	0,08	0,98	0,99	0,17	0,36	1,32	0,86	0,16	0,29	1,00
12 Fsup_4	0,88	0,09	0,65	0,98	1,00	0,02	0,94	1,06	0,96	0,04	0,84	0,99
13 Fsup_5	0,88	0,07	0,73	0,98	1,00	0,02	0,96	1,05	0,96	0,03	0,89	0,99

5.2 *SAE Model 2*

This concerns EBLUP models for going from NUTS2 to NUTS1 level, again utilising in combination survey data and the information compiled in NewCronos.

The list of target variables and covariates is the same as that for SAE Model 1 described above. However, there are differences in the extent to which the variables are available and the nature of the countries involved in the same model. Note that because of the lack of NUTS2 identifiers in the microdata, SAE Model 2 is based only on pooled information from five countries. Table 4 shows some performance measures of SAE Model 2. For each model (target variable), three measures are shown as in Table 3.

The performance of the model in terms of gain in efficiency is obviously better for Model 2 (NUTS2 level) compared to Model 1 (NUTS1 level). This is because here the sample sizes available for direct estimates are smaller. The highest gains, of 20-25%, are for LAT-C, MAN-c and FSUP-1 deprivation measures. Again, as with Model 1, the gain for HCR_N2 is almost twice as large as that for HCR_C. This is important in the context of constructing regional indicators. The gain for HCR_C, FSUP-2, FSUP-4 and FSUP-5 is around 10%, while no prediction is possible for FSUP-3 because of the lack of adequate data. For the logarithm of equivalised income and the logarithm of the per capita income, the relative gains are the smallest among the variables.

Table 4 – *Performance measures for the SAE Model 2***.**

	Gamma				Estimate EBLUP/direct estimate				Standard error (SE) SE(EBLUP)/SE(direct estimate)			
	mean	CV	min	max	mean	CV	min	max	mean	CV	min	max
1 HCR_C	0,80	0,22	0,45	0,98	1,01	0,08	0,86	1,34	0,90	0,11	0,71	1,00
2 HCR_N2	0,66	0,38	0,19	0,95	1,01	0,07	0,83	1,30	0,82	0,22	0,47	1,00
3 logEqInc	0,81	0,23	0,44	0,98	1,00	0,00	1,00	1,01	0,94	0,18	0,68	1,35
4 logIncPC	0,85	0,14	0,65	0,99	1,00	0,00	0,99	1,01	0,92	0,12	0,74	1,21
5 FM_C	0,75	0,27	0,40	0,98	1,02	0,14	0,80	1,63	0,88	0,12	0,66	1,02
6 FS_C	0,68	0,32	0,32	0,97	1,02	0,09	0,85	1,45	0,85	0,14	0,63	0,99
7 Latent	0,61	0,36	0,23	0,96	1,01	0,08	0,84	1,41	0,81	0,16	0,50	0,98
8 Manifest	0,55	0,49	0,12	0,97	1,06	0,25	0,71	2,25	0,76	0,24	0,36	1,00
9 Fsup_1	0,60	0,41	0,22	0,97	1,01	0,08	0,86	1,28	0,80	0,18	0,54	1,00
10 Fsup_2	0,73	0,22	0,47	0,97	1,01	0,07	0,87	1,26	0,88	0,09	0,70	0,99
11 Fsup_3												
12 Fsup_4	0,77	0,15	0,51	0,97	1,01	0,05	0,88	1,24	0,90	0,06	0,76	0,99
13 Fsup_5	0,76	0,22	0,49	0,98	1,00	0,05	0,87	1,11	0,89	0,10	0,72	1,01

6. Illustrative results

6.1 *Poverty rates with country and region-specific poverty lines*

In order to understand the effect of using poverty lines defined at different levels, it is useful to begin with regional differences in mean income.

Figure 1 reports the mean net equivalent income at NUTS2 level. The New Member States have much lower income levels (in PPS) compared to former EU15 countries. More directly relevant are the large regional disparities within countries like Italy and Spain. Note also the higher mean incomes in metropolitan centres (Paris, London).

Now turning to estimates of regional poverty rates (conventional HCR), it should be noted that the results are based on an extremely large number of computations, performed separately for each survey wave in each country. These are then used to construct consolidated measures of the type we recommend for regional analysis (Section 4.1).

Figure 2 shows the concentration of the areas with the highest poverty rates (bottom quintile) to be in Portugal, Spain, Greece and Southern Italy. The highest estimated poverty rate using country poverty lines is in Sicilia (with 39.4%), and the next highest in Calabria (with 39.2%). In the UK as well, the proportion in poverty is also quite high in many areas outside the South East. In Italy it is interesting to note the striking difference between the South and the North.

Figure 1 – *Net equivalent income – NUTS2.*

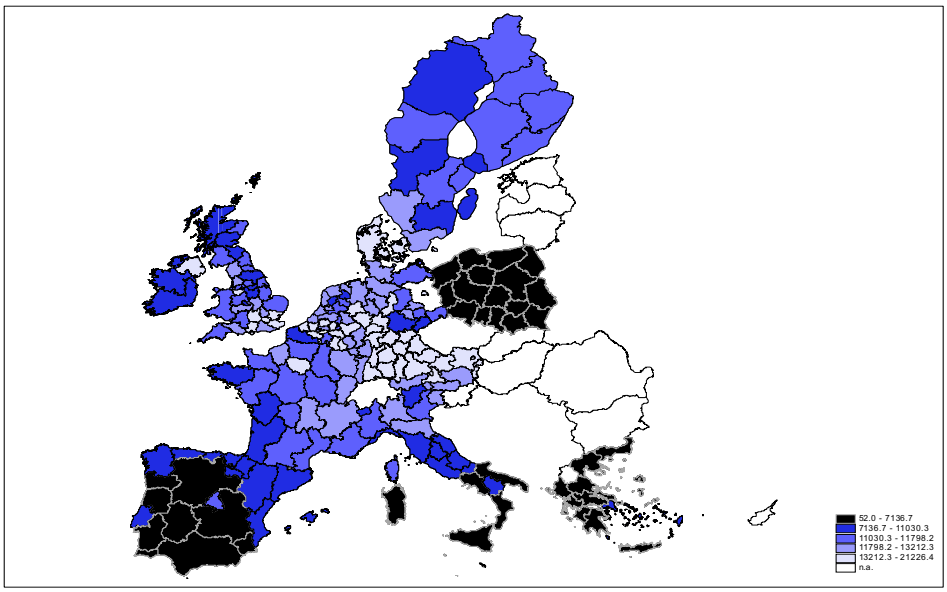
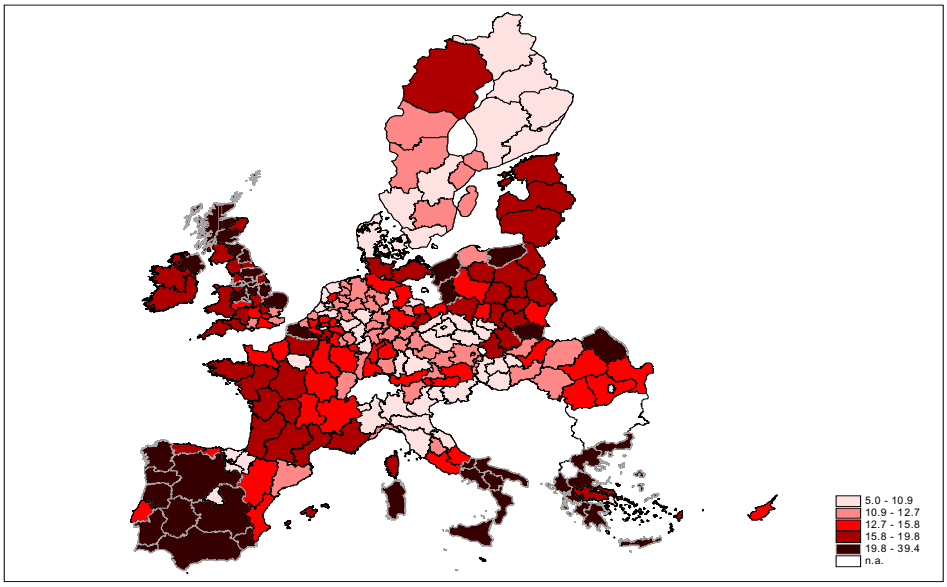


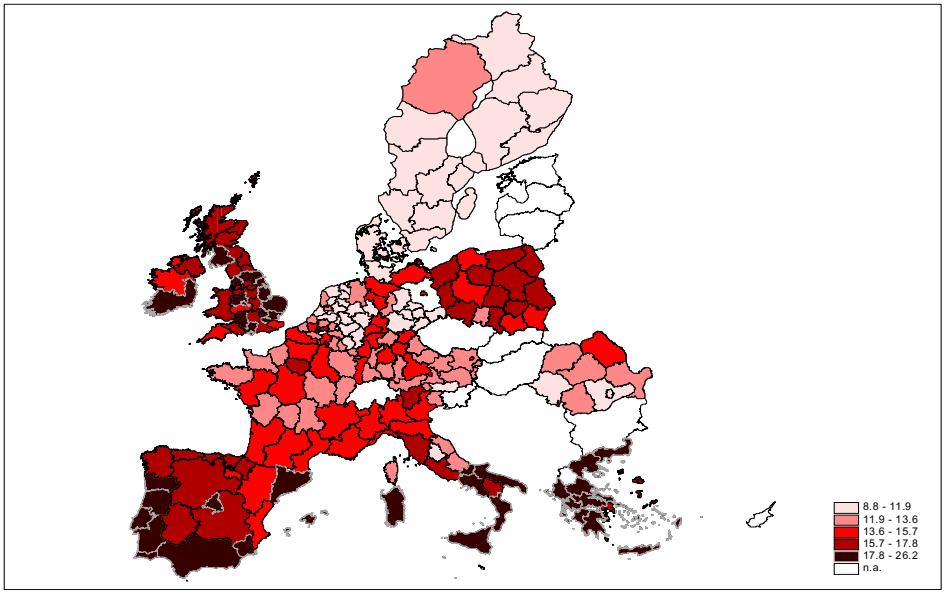
Figure 2 – *Head Count Ratio NUTS2 regions (country poverty lines).*



In Figure 3 with NUTS2 poverty lines, we see a less homogenous situation in terms of head count ratio. It is the effect of the definition of the poverty line. Defining the poverty line at NUTS2 level makes the poverty measure more truly relative. It is interesting to note that some areas that show the worst situation in Figure 2 do not belong to the bottom category in Figure 3 (for example Basilicata in Italy). On the other hand, some areas that belong to the best class in Figure 2, move to the middle bracket in Figure 3, such as Toscana, Emilia Romagna and Lombardia in Italy. The same applies to some regions of Spain.

Countries where regional differences in levels of income are small tend to present similar pictures irrespective of whether the country or NUTS2 poverty lines have been used.

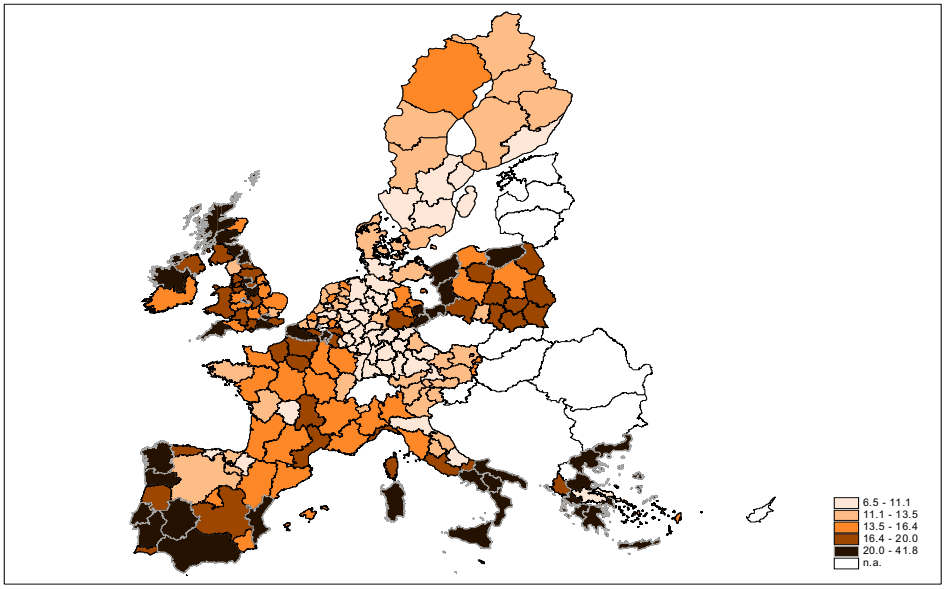
Figure 3 – *Head Count Ratio NUTS2 regions (NUTS2 poverty lines).*



6.2. Non-monetary deprivation

Figure 4 shows the variation of the overall non-monetary deprivation rate across NUTS2 regions. We observe very high values of deprivation in Portugal, South West Spain and South Italy. A better off country is Germany, showing a very homogenous behaviour among regions.

Figure 4 – *Overall non-monetary deprivation rates (NUTS2 regions).*

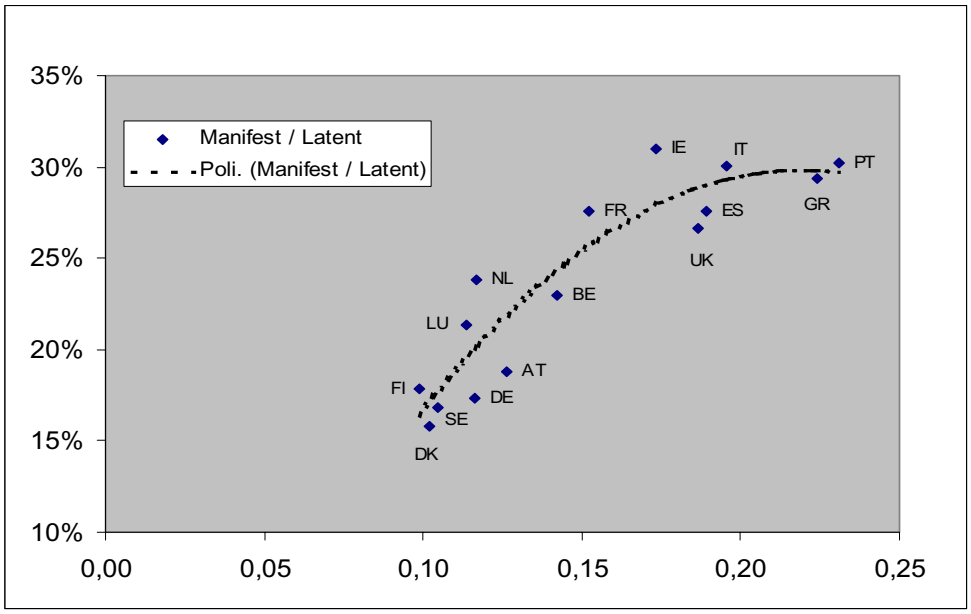


The objective of Figure 5 is to display the degree of overlap, at the level of individual persons, between monetary and non-monetary forms of deprivation. The figure shows the manifest deprivation index as a percentage of the latent: it can be interpreted as an index of the degree of overlap (or intersection), at the level of individual persons, between income poverty and non-monetary deprivation.

By definition, this ratio varies from 0 to 1. When there is no overlap (i.e., when the subpopulation subject to income poverty is entirely different from the subpopulation subject to non-monetary deprivation), manifest deprivation rate and hence the above mentioned ratio equals 0. When there is complete overlap (i.e., when exactly the same subpopulation is subject to both income poverty and non-monetary deprivation), the manifest and latent deprivation rates are the same and hence the above mentioned ratio equals 1.

It is important to highlight that there is a higher degree of overlap between income poverty and non-monetary deprivation at the level of individual persons in poorer countries, and a lower degree of overlap in richer countries. In richer countries different dimensions have less overlap, and hence deprivation appears to be more multidimensional. In poorer countries, there is more overlap, making deprivation more intensive for the individuals involved.

Figure 5 – *Manifest deprivation rate as a percentage of latent deprivation rate, against a measure of the level of poverty or deprivation in the country. EU15.*



Acknowledgments

This paper provides an overview of a methodology and application for the construction of regional (sub-national) indicators of poverty and social exclusion in Europe. It is based on research carried out under the project Regional Indicators to Reflect Social Exclusion and Poverty of the European Commission, Employment and Social Affairs DG. The project was carried out by Centro di Ricerca Interdipartimentale sulla Distribuzione del Reddito "C. Dagum" (C.R.I.DI.RE.) at Università degli Studi di Siena. Achille Lemmi acted as the Project Director, and Vijay Verma as the Research Director for the project. Other main contributors to

the research included Gianni Betti, Anna Mulas, Michela Natilli, Laura Neri, and Nicola Salvati.

The empirical results and the approach of the paper has also represented the starting point and scientific benchmark for the Italian PRIN research project n. 2007HEWTBE_003 coordinated at National level by Prof. G.M. Giorgi and locally carried out by the above mentioned C.R.I.DI.RE. research centre.

References

Battese G.E., Harter R.M., Fuller W.A. (1988), An error-component model for prediction of county crop areas using survey and satellite data, *Journal of the American Statistical Association*, **83**, pp. 28-36.

Betti G., Verma V. (2008), Fuzzy measures of the incidence of relative poverty and deprivation: a multi-dimensional perspective, *Statistical Methods and Applications*, **12(2)**, pp. 225-250.

Fay R.E., Herriot R.A. (1979), Estimates of income for small places: an application of James-Stein procedure to census data, *Journal of the American Statistical Association*, **74**, pp. 269-277.

Giorgi L, Verma V. (2002), European Social Statistics: Income, Poverty and Social Exclusion: 2nd Report. Luxembourg: Office for Official Publications of the European Communities.

Gosh M., Rao J.N.K (1994), Small Area Estimation: An Appraisal (with discussion), *Statistical Science*, **9 (1)**, pp. 55-93.

Handerson C.R. (1950), Estimation of Genetic Parameters, *Annals of Mathematical Statistics*, **21**, pp. 309-310.

Lemmi A., Betti G. (2006), *Fuzzy Set Approach to Multidimensional Poverty Measurement*. New York: Springer.

Rao J.N.K (2003), *Small Area Estimation*. Wiley, London.

Verma V., Betti G., Lemmi A., Mulas A., Natilli M., Neri L., Salvati N. (2005), Regional indicators to reflect social exclusion and poverty. Final report. Project VT/2003/45, European Commission, Employment and Social Affairs D.G.

SUMMARY

This paper provides an overview of a methodology and application for the construction of regional (sub-national) indicators of poverty and social exclusion in Europe. Starting from Laeken Indicators defined at the country level, proposals are made for their regional adaptation. A strategy is developed and implemented for the construction of regional indicators. It involves making use of micro-level survey data and region-level data from other sources to produce regional indicators using composite area-level EBLUP methodology. Illustrative results are shown of estimates down to NUTS2 level for most EU regions.

Gianni BETTI is Associate Professor of Economic Statistics at the Department of Quantitative Methods of the University of Siena, and member of the C.R.I.DI.RE. research centre (Centro Interdipartimentale di Ricerca sulla Distribuzione del Reddito "C. Dagum").

Achille LEMMI is Full Professor of Economic Statistics at the Department of Quantitative Methods of the University of Siena, and Director of the C.R.I.DI.RE. research centre (Centro Interdipartimentale di Ricerca sulla Distribuzione del Reddito "C. Dagum").

Vijay VERMA is Research Professor of Social Statistics at the Department of Quantitative Methods of the University of Siena, and member of the C.R.I.DI.RE. research centre (Centro Interdipartimentale di Ricerca sulla Distribuzione del Reddito "C. Dagum").

Rivista Italiana di Economia Demografia e Statistica Volume LXIII nn. 1-2 – Gennaio-Giugno 2009

ANALYSIS OF GINI FOR EVALUATING ATTRITION IN ITALIAN SURVEY ON INCOME AND LIVING CONDITION

Claudio Ceccarelli, Giovanni Maria Giorgi

1. Introduction

European Community Statistics on Income and Living Conditions (Eu-Silc) is a set of statistical indicators of income, poverty and social exclusion which has been regulated by the European Parliament since 2003. In particular, the Regulation defines the responsibilities of Member States and Eurostat and lays down a set of rules to improve data quality, comparability and timeliness besides promoting a better integration of new surveys within national statistical systems.

In order to comply with all tasks entrusted by Eurostat and to deepen the analysis of income distribution, living conditions, inequality and poverty in Italy, the Italian National Institute of Statistics (Istat) has set up a survey on income and living condition (hereinafter It-Silc), which is substantially made of a cross-sectional and a longitudinal component.

Several studies show how the phenomenon of *selective attrition* may create a bias in the evaluation of results of analysis carried out by panel surveys due to non-random mechanisms generating non response from wave to wave.

According to Rendtel (2002, p.4), panel attrition "is defined by unit non-response of eligible persons or households that occurs after the first wave of panel".

Our paper aims at proving, by a decomposition of Gini concentration index, also known as Analysis of Gini (ANOGI) (see Frick et al., 2006), whether attrition introduces an element of bias in the analysis of income distribution at the regional level.

The work is organized as follows: in section 2, attrition is briefly described along with the main causes which may generate it in statistical surveys on households and individuals. In section 3, It-Silc main features are highlighted; section 4 gives a description of the methodology underlying the ANOGI; section 5 focuses on attrition patterns and after showing main results (section 6), the work ends with some concluding remarks and possible future developments (section 7).

2. Attrition

The substantial difference between repeated longitudinal and cross-sectional surveys lies in the sample of statistical units analyzed from time to time in the survey. In cross-sectional surveys, indeed, the initial theoretical sample is randomly drawn from population registers and is generally fixed, excluding those variations introduced during the survey design phase. In longitudinal surveys, from the second wave onwards, the sample size varies in function of previous wave respondents and of the different characteristics of the survey design which sets the mechanism generating the theoretical sample of the following wave. Within this context, the phenomenon of non-response assumes different connotations depending on whether it emerges during the first or the following waves.

From a merely theoretical and conceptual point of view, non-response at the first wave takes on characteristics which are absolutely similar to what happens in cross-sectional surveys. From the second wave onwards, non-response assumes different connotations. It may happen, indeed, that some statistical units, after responding to the first wave, choose not to participate in the survey anymore and drop out of the sample, albeit they are still eligible units: this is what is called attrition.

Attrition can be caused by different reasons which can be related either to fieldwork or to response behavior. Any mistake or lack which can inevitably occur during the various operational micro-phases of the survey may result in non participation. Among the others, it is worth mentioning the incapability to trace respondents with a high degree of mobility throughout the territory; non correct application of rules for the conduction of surveys; changes in survey technique or in questionnaire, which may induce refusal to participate; incorrect implementation of rules to trace sample units throughout the territory. Some additional causes, more or less subjective and depending on respondents' behavior and interaction with survey operators, can result in refusal to continue to participate in the panel. Impossibility to participate for health reasons or diffidence caused by change in interviewer from a wave to another or unconditioned refusal are some examples.

In panel surveys, sample units decrease in function of demographic exits due to individuals' death and migrations. Such units are no longer eligible units as they represent a part of population who is no longer considered as benchmark but defined out of scope.

There are also some cases in which attritors start participating again. Such participation pattern is called "temporary drop-out" and is due, for instance, to temporary impossibility to participate in the survey or to simple change of mind.

3. Italian survey on income and living condition (It-Silc)

The It-Silc is defined within the European Regulation no. 1177/2003 which outlines its main methodological, thematic and organizational aspects. In order to ensure the comparability of data collected by all Member States, common rules have been set for the following themes: sampling and tracing, definitions, list of primary variables, fieldwork aspect and imputation procedures, intermediate and final quality reports.

3.1 *Sample design*

The sample design planned and implemented in function of the main estimates which the survey has to produce and the planned study domains, is based on four independent longitudinal samples. Such design, called rotation design, provides that every year the longitudinal sample be closed after reaching the fourth wave and a new sample be started.

Each longitudinal sample is a two-stage sample: the primary sample units, municipalities, are stratified by region and demographic size, while the secondary sample units, households, are drawn from the population register of sampled municipalities.

Table 1 – *Rotational sample scheme in It-Silc.*

Samples	Years						
	2004	2005	2006	2007	2008	2009	2010
c1	$W_{(4)}$						
c2	$W_{(3)}$	$W_{(4)}$					
c3	$W_{(2)}$	$W_{(3)}$	$W_{(4)}$				
c4	W_1	W_2	W_3	W_4			
c5		W_1	W_2	W_3	W_4		
c6			W_1	W_2	W_3	W_4	
c7				W_1	W_2	W_3	W_4

In 2004, the first four longitudinal samples (*c1, c2, c3* and *c4* in Table 1) participate in the survey all for the first time. To start rotation, *c1* sample is assumed to be at its fourth and last wave ($W_{(4)}$ in Table 1), *c2* sample at its third wave ($W_{(3)}$), *c3* sample at its second wave ($W_{(2)}$) and *c4* sample at its first wave (W_1). The sample *c4* is the first longitudinal one which, started in 2004, will go on correctly over four waves, as per design, and in 2007 will allow the realization of

the first complete longitudinal sample (made up of W_1, W_2, W_3, W_4). The new longitudinal sample *c5* starts in 2005 and takes the place of *c1*, dropped in 2004. Generally, a new longitudinal sample is made up of the same first-phase units (municipalities) and new second-phase units (households).

The cross-sectional sample results every year from the union of the four longitudinal samples, each one for its specific wave: thus, each cross-sectional sample includes one fourth of households participating in the survey for the first time, one fourth of households participating for the second time, one fourth for the third time and one fourth for the fourth time.

The initial cross-sectional sample, relating to year 2004, is made up of about 32,000 households in all, that is 8,000 for each longitudinal sample.

For the year 2005, the cross-sectional sample size is given by the sum of the following items:

- number of households with individuals responding in the first wave for longitudinal samples *c2*, *c3* and *c4*;
- 8,000 newly drawn households belonging to the new longitudinal sample *c5*.

In this way, a household that has not been drawn for the first wave can enter the sample if joined with one sample member dropped out from the origin household.

The same procedure is used for the sample determination over the following years.

In the hypothesis of simple random sampling and given a level of sampling error, Eurostat fixes the minimum sample size; the definition of the sample size to be realized, according to which the whole survey is planned, comes from the hypothesis on design effect related to the sampling designs carried out by the various Statistical Agencies as well as from the supposed response rates by the survey. In longitudinal surveys the assessment of response rates requires, moreover, specification of an attrition trend.

3.2 *Cross-sectional weighting*

The cross-sectional weighting strategy develops through the following phases, which are usually used for the construction of estimators in various Istat's social surveys:

1) definition of design weight as the inverse of inclusion probability;
2) calculation of coefficients of correction for non-response bias;
3) determination of final cross-sectional weighting adjusted on according to known totals derived from external data relating to the distribution of households and persons in the target population.

The design weight is directly derived from the sampling design.

The second step is based on the hypothesis that the process generating non-response is not missing at random mechanism. In this case, a strategy is applied which follows the same criteria as weighting cells in order to single out sub-populations in which equal response behaviour may be assumed among those who have participated in the survey and those who have not. Sample households have been partitioned into cells through segmentation obtained using a *Chaid*-based decision tree (*Chi-squared automatic interaction detection*; Kass, 1980). Such a method consists in splitting the sample in sub-groups according to the relationship between ratio of response rate and explicative variables. The methodology underlying the *weighting cells* belongs to group of explicit modeling techniques to reduce non-response bias. In It-Silc, making use of both personal and fiscal data already available during the definition of the sample, a partition into homogeneous cells has been obtained in which is possible to adopt the hypothesis of missing at random non-response mechanism. To realize the cells, following data have been used: demographic size of the municipality; citizenship of the reference individual; region of residence; distribution of households by number of components; distribution of households by income group.

In order to calculate final cross-sectional weights, calibration estimators (see Deville and Särndal, 1992) are used. As provided for by Eurostat, each longitudinal sample at the first wave is bound to: resident population by geographical area, sex and age class[1], income reference year (31st December of year *t-1*), number of resident households by region as on date of survey (31st December of year *t*).

A brief mention should also be made of determination of final weight of the cross-sectional part of the survey for the years following the first. The weighting procedure, indeed, has to take into account that the cross-sectional sample comprises a longitudinal sample (e.g., *c5* for the 2005 sample) of households who have participated in the first interview and three samples (*c2*, *c3*, *c4*) who are already at their second interview. From a methodological point of view, the inclusion probability for households with sample individuals changes: three-fourths of the sample, indeed, who are not present in the first wave are no longer likely to be included for the calculation of design coefficient. For the households of individuals belonging to these three samples, however, the *weight share method* is used as if it were the design weight. The weight is defined as follows (see Istat, 2008):

[1] Age classes are: 0-15, 16-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-74, 75 and over.

$$\omega_h^t = \frac{\sum_{i \in s_0^t \cap h} \psi_i^t}{N_h^t} \tag{1}$$

where s^t is the entire sample, including new entrants, and $s_0^t \subset s^t$ is the longitudinal sample (individuals aged 14 and over belonging to the sample in the first year of the panel) of respondents in year *t*. ψ_i^t is the initial weight of the individual *i* in the year *t*, calculated as described above. N_h^t is the total number of sample and non-sample members of the household *h*. By construction, this household weight takes account of the correction for non-participation in the years following the first. Subsequently, after correcting non-response bias in the only new entrant sample *c5*, the whole cross-section undergoes the calibration procedure as described above.

3.3 *Longitudinal weighting*

A longitudinal sample produces estimates referred to the target population of the same year when the sample first participates in the survey (see Osier *et al.,* 2006). The longitudinal population in year *t+1*, includes individuals of the population in year *t* and excludes drop-outs between the year *t* and year *t+1* (OUT_{t+1}).

The target longitudinal population started in 2004 covering the years 2005 and 2006 assumes the following form:

$$P_{2005}^{(L)} = P_{2004} - OUT_{2005} \tag{2}$$

$$P_{2006}^{(L)} = P_{2005}^{(L)} - OUT_{2006} = P_{2004} - OUT_{2005} - OUT_{2006} \ . \tag{3}$$

In general terms, given a panel in year *n* starting in year *t=1*, the longitudinal target population is equal to:

$$P_n^{(L)} = P_1 - \sum_{t=2}^{n} OUT_t \ . \tag{4}$$

It is to be noted that the longitudinal population at time *t* ($P_t^{(L)}$) differs from the population at time *t* (P_t) in that it does not include individuals born or migrated into the reference population at time *t=1*.

Table 2 shows the values of the longitudinal population in respect of the whole population and people aged 16 and over in function of the various samples and respective waves. In year 2005, sample *c5* refers to the resident population in Italy on 31st December 2004, while sample *c4* refers to resident population in Italy on 1st January 2004 net of drop-outs (deaths and migrations) during the year 2004.

The strategy adopted for the weighting procedure develops through the following phases: determination of the design weight, calculation of attrition correction coefficients and determination of final weight (base weight). In the first wave, the theoretical sample drawn from municipal population registers, along with its design weight provides an estimate of the resident population.

Table 2 – *It-Silc longitudinal population by sub-samples and years (thousand).*

Sub-sample		ANNI		
		2004	2005	2006
c3	16 years and over	48,762	48,554	48,113
	Total	57,952	57,266	56,578
c4	16 years and over	48,762	48,554	48,113
	Total	57,952	57,287	56,594
c5	16 years and over	-	49,286	48,762
	Total	-	58,418	57,712

The "theoretical" sample to be interviewed in the second wave is composed of first wave respondents; coupled with its final weights (calculated as described in sub-section 3.2), it represents the resident population in year *t*.

Formally, let *c4* be the sample starting its longitudinal path in 2004 and Ψ_{2004} the vector of final weights of the year 2004. By construction, these weights make sample *c4* representative of the resident population in the year 2004.

$$(c4_{2004}, \Psi_{2004}) \rightarrow P_{2004} \ . \tag{5}$$

At time *t+1* (i.e. 2005) the longitudinal sample includes the initial sample net of drop-out from the reference population (out_{2005}) and those who do not participate in the survey, albeit still eligible (x_{2005})[2]:

$$c4_{2005}^{(L)} = c4_{2004} - out_{2005} - x_{2005} \ . \tag{6}$$

Assuming that sample drop-outs (out_{2005}), weighted by Ψ_{2004}, are an estimate of population drop-outs (OUT_{2005}), we have[3]:

$$\left(\left(c4_{2005}^{(L)} + x_{2005}\right), \Psi_{2004}\right) \rightarrow \left(P_{2004} - OUT_{2005}\right) \ . \tag{7}$$

In general terms, given a generic sample *j,* started in the year n_o, we have in the year *n*

$$\left(\left(cj_n^{(L)} + \sum_{t=n_0+1}^{n} x_t\right), \Psi_{n_0}\right) \rightarrow \left(P_{n_0} - \sum_{t=n_0+1}^{n} OUT_t\right) \ . \tag{8}$$

For the remaining sample to be representative of the initial population net of reference population drop-outs, weights have to be changed so as to take into account eligible units who stop participating in the survey. It follows that

[2] The longitudinal sample is also net of a third group of individuals, namely those who do not participate in the survey and, due to lack of information about them, it is not even clear whether they are still part of the target population or not. Therefore, each individual is to be assigned to one group or the other. In It-Silc surveys, a logistic regression model is used to estimate the propensity to stay in the population as a function of a set of explicative variables (age is obviously the most influential variable). This model is initially applied to the group of sample individuals for whom necessary information is available; the same parameters are then applied to individuals for whom no information is available for determining to which group to assign them.

[3] It is to be noted that such information cannot be obtained from external sources. Istat's demographic balance, indeed, gives the total of deaths and migrants (leaving out those who move into cohabitations, who are less numerous) but does not specify whether they belonged to the initial population. For instance, should an individual die during the period in question we could never know if he/she has just entered (by immigration, birth or following a move from an institution) or are already part of the target population.

$$\left(cj_n^{(L)}, \Psi_n\right) \rightarrow \left(P_{n_0} - \sum_{t=n_0+1}^{n} OUT_t \right) \tag{9}$$

where $\left(cj_n^{(L)}, \Psi_n\right)$ represents the initial population excluding drop-outs between n_0 and n.

The non-randomness of non-participation in the survey introduces an element of bias in estimates of aggregates. The basic idea to correct weights year after year was to work on the sample at the first wave in order to make it as representative of the initial population as possible, while taking into account those individuals who are still in the sample and inflating their weights in consideration of non randomness of attrition. From a practical point of view, an updating process has been set up which, starting from the individual weight of the sample unit and considering the various factors accounting for individuals remaining in a panel, leads to new individual weights. The method applied to inflate weights is the segmentation of the sample in homogeneous cells; the same as that described for the cross-sectional surveys with the only difference that in longitudinal surveys it is possible to use all information on individuals collected in previous years in order to determine the best partition possible. The so-called *base weight* is thus obtained, which is different over years even when it refers to the same individual, and by which it is possible to carry out longitudinal analyses: a real longitudinal weight, indeed, does not exist, but can be calculated starting from the base weight relating to individuals remained in the sample for the duration of the panel, but this basically depends on the nature of the analysis to carry out (see Osier *et.al.,* 2006; Ceccarelli and Cutillo, 2007).

4. Analysis of Gini (ANOGI): some methodological remarks

Among the indices used in the literature to investigate, for instance, the inequality of the income distribution, the Gini (1914) concentration ratio has again a role of primary and fundamental importance. Sometimes researchers have proposed different formulas from the original one with the purpose to fully exploit the application potentialities of the Gini index (G) in the most disparate fields[4].

[4] For a survey of the topical interest, new interpretations and extensions of the Gini index, see, e.g., Giorgi (1990, 1992, 1993, 1999, 2005).

In this context, in order to derive a useful decomposition by population subgroups, Lerman and Yitzhaki (1984) expressed G in terms of covariance[5] between a variable y (e.g. income) and its cumulative distribution function $F(y)$, that is:

$$G = \frac{2}{\mu}\mathrm{cov}\left[y, F(y)\right] \tag{10}$$

where μ is the mean of y.

Now, let us consider a population (P) divided in k subpopulations or groups $P = P_1 \cup P_2 \cup ... \cup P_k$ the Gini index can be written as (Yitzhaki, 1994, p.154)

$$G_u = \sum_{i=1}^{k} s_i G_i O_i + G_b \tag{11}$$

where

$s_i = p_i \bar{y}_i / \mu_u$ is the ratio between the mean of variable y in the subpopulation i weighted by the units presents in it (p_i) and the mean of y calculated on the whole population;
G_i is the Gini index within subpopulation i;
O_i is the *overlapping index* of subpopulation i with the entire population;
G_b is the between-subpopulations inequality.

The analysis based on formula (11) is known as Analysis of Gini or ANOGI (see Frick *et al*., 2006), that is similar to the Analysis of variance (ANOVA). In particular, the overlapping index is the element which conceptually distinguishes ANOGI and ANOVA, while G_b, albeit it can be negative (see Yitzhaki and Lerman, 1991, p. 322, note 9), is similar in its meaning to the between-group variance of ANOVA, i.e. it indicates the degree of inequality between subpopulations in terms of concentration.

4.1 *Overlapping index*

A brief mention should also be made of problems related to stratification and overlapping in the analysis of distribution of some variables (e.g income). Generally speaking, there is *stratification* (see, Yitzhaki, 1988, p.39; Yitzhaki and

[5] See also De Vergottini (1950, p. 453), Stuart (1954), and Piesch (1975, p. 39).

Lerman, 1991, p.319) when a group is isolated from the members of other groups. More specifically, there is perfect stratification when the members of a group occupy distinct range within an overall distribution and no member of a group is included in the same range of another group. A classical example is the subdivision of a population into income deciles. Each unit of a given decile belongs exactly to the range of the considered decile. In the absence of stratification, overlapping occurs.

In brief, being $\overline{F}_{ui}(y) = \int F_u(y) dF_i(y)$ the expected rank of the units belonging to group *i* within the distribution the entire population, and given that quantity $\text{cov}_i(y, F_u(y)) = \int (y - \mu_i)(F_u(y) - \overline{F}_{ui}(y)) f_i(y) dy$ represents the covariance between *y* (income) and rank of units belonging to group *i ,* calculated on their position in the overall distribution, the overlapping index O_i may be expressed as:

$$O_i = O_{ui} = \frac{\text{cov}_i(y, F_u(y))}{\text{cov}_i(y, F_i(y))} \ . \tag{12}$$

In this case, the index measures the degree of overlapping between the distribution of the units belonging to group *i* with the distribution of the entire population. In other words, there is perfect stratification (as in the case of income deciles) when the units of the group *i* occupy the same relative position both in the population and in the group distribution.

With reference to the population partition in *k* groups, the overlapping index referred to a given group *i* is expressed by the following formula:

$$O_i = \sum_j p_j O_{ji} = p_i O_{ii} + \sum_{j \neq i} p_j O_{ji} = p_i + \sum_{j \neq i} p_j O_{ji} \tag{13}$$

where (Yitzhaki, 1994)

$$O_{ji} = \frac{\text{cov}_i(y, F_j(y))}{\text{cov}_i(y, F_i(y))} \ . \tag{14}$$

The formula (14) represents the overlapping index of group *j* by group *i* and provides a measure of the presence of group *j* units within the group *i*. In particular, the main properties are (Frick *et al*., 2006, p.437):

i. O_{ji}=0, no member of the *j* group lies in the range of distribution *i.* Group *i,* therefore, is a "perfect stratum", i.e. its range is not "contaminated" by members of the *j* group.
ii. O_{ji}=1, the distributions of group *i* and *j* are identical, being $O_{ii} = 1$.
iii. O_{ji} is not symmetrical, that is the higher O_{ji} the lower O_{ij}.
iv. $O_{ji} \leq 2$. If all observations of distribution *j* are in the range of *i* and are concentrated at the mean of distribution *i* then O_{ji} assumes the maximum value (Yitzhaki 1994, p.151).

4.2 *Between-group inequality (G_b)*

Another essential element of ANOGI is the measurement of the between-group inequality *(G_b)* defined as:

$$G_b = \frac{2\,\mathrm{cov}(\mu_i, \overline{F}_{ui}(y))}{\mu_u} \tag{15}$$

which is twice the covariance between the mean of variable *y* of each group and the groups' mean rank in the whole population, divided by the mean of *y* calculated on the whole population.

Pyatt (1976) introduced a type of between-group inequality measure (G_b^p) based on the assumption of perfect stratification. In this case, the covariance is calculated between the mean of each group and the groups' mean rank. From a strictly formal point of view, this is defined as:

$$G_b^p = \frac{2\,\mathrm{cov}(\mu_i, \overline{F}_i(y))}{\mu_u} \ . \tag{16}$$

From a conceptual point of view, it may be argued that G_b is not really a concentration index because, as mentioned earlier, it can be negative. As per formula (11), moreover, in case of perfect stratification – overlapping index equals to zero – the G_b indicator reaches its upper level as the quantification of the amount of total inequality is explained by between-group inequality.

It derives that (see Yitzhaki and Lerman, 1991, p. 322)

$$G_b^p \geq G_b \ . \tag{17}$$

With simple algebra (11) can be written as

$$\begin{aligned} G_u &= \sum_{i=1}^{k} s_i G_i + \sum_{i=1}^{k} s_i G_i (O_i - 1) + G_b^p + (G_b - G_b^p) = \\ &= IG + IGO + BG + BGO \end{aligned} \tag{18}$$

and these four components of ANOGI may be conceptually compared to ANOVA.

Frick *et al*. (2006, p.438-440) schematize the comparison between ANOGI and ANOVA as follow:

Component similar to ANOVA

Within $\quad IG = \sum_{i=1}^{k} s_i G_i \quad 0 \leq IG \leq G_u$

Between-Pyatt $\quad BG = G_b^p \quad 0 \leq BG \leq G_u$

Additional component respect to ANOVA

Within $\quad IGO = \sum_{i=1}^{k} s_i G_i (O_i - 1)$

Between $\quad BGO = (G_b - G_b^p) \quad -BG - IGO - IG \leq BGO \leq 0$.

5. Response pattern

On the basis of definitions given in section 2, three groups of individuals have been identified: respondents, those namely who have actively participated in a survey; *out of scope,* all individuals who, after responding to the previous wave*,* exit from the target population (moved abroad, moved to institutional household, deaths); attritors*,* all individuals, that for various reasons, have not participated in the survey even after responding to the previous wave*,* excluding those who exited from the sample and entered it again. These latter have been excluded as they are

not representative of a monotonic response pattern. In terms of response behaviour, those who have second thoughts cannot have the same approach to the survey as those who definitely exit it.

Table 4 – *Response pattern for wave and geographical area.*

Geographical area	Waves						
	2004	2005			2006		
North-West	Respondent	*Sample size*	*6,385*	*100.0*			
		Out of scope	117	1.8			
		Attritors	882	13.8			
		Respondent	5,386	84.4	*Sample size*	*5,386*	*100.0*
					Out of scope	52	1.0
					Attritors	805	14.9
					Respondent	4,529	84.1
North-East	Respondent	*Sample size*	*6,248*	*100.0*			
		Out of scope	96	1.5			
		Attritors	774	12.4			
		Respondent	5,378	86.1	*Sample size*	*5,378*	*100.0*
					Out of scope	78	1.5
					Attritors	595	11.1
					Respondent	4,705	87.4
Centre	Respondent	*Sample size*	*6,301*	*100.0*			
		Out of scope	104	1.7			
		Attritors	952	15.1			
		Respondent	5,245	83.2	*Sample size*	*5,245*	*100.0*
					Out of scope	61	1.2
					Attritors	588	11.2
					Respondent	4,596	87.6
South	Respondent	*Sample size*	*5,271*	*100.0*			
		Out of scope	66	1.3			
		Attritors	395	7.5			
		Respondent	4,810	91.2	*Sample size*	*4,810*	*100.0*
					Out of scope	50	1.0
					Attritors	319	6.6
					Respondent	4,441	92.4
Islands	Respondent	*Sample size*	*2,130*	*100.0*			
		Out of scope	46	2.2			
		Attritors	213	10.0			
		Respondent	1,871	87.8	*Sample size*	*1,871*	*100.0*
					Out of scope	13	0.7
					Attritors	117	6.3
					Respondent	1,741	93.0
Italy	Respondent	*Sample size*	*26,335*	*100.0*			
		Out of scope	429	1.6			
		Attritors	3,216	12.2			
		Respondent	22,690	86.2	*Sample size*	*22,690*	*100.0*
					Out of scope	254	1.1
					Attritors	2,424	10.7
					Respondent	20,012	88.2

The longitudinal component in It-Silc survey shows how the various types of non-response depend on different design characteristics. In Table 4, for instance, the response pattern of samples c3 and c4 (those who have reached the third wave) are analyzed in order to evaluate how the length of panel may affect attrition.

Table 4 shows the response patterns in Italy and by geographical area; in particular, it focus on the very low percentage of out of scope individuals (oscillating around 1.6%). It is, indeed, to be noted that the said percentage regarding both the second and the third wave is quite stable, with the only exception of the Islands, where it varies between 2.2% and 0.7%. The above values prove that the It-Silc sample is adequate to represent the drop-outs from the longitudinal population regardless of the wave.

The most interesting element is the trend of attritors. The total sample shows an attrition level after one year equal to 12.2%, which reduces to 10.7% after two years. With the only exception of the North-West, where the attrition level rises from 13.8% to 14.9%, the same downward trend is reported in other geographical areas, especially in the Islands.

Respondents' different behaviours may be accounted for by various factors, such as, for instance, the different structure of interviewers' network, or citizens' different awareness of the importance of official statistics, or the well-known difficulties to obtain high response rates in large cities. From the analysis of the pattern of those who have responded to all three waves, indeed, it emerges that the South (84.3%) has a higher rate of permanence in the sample than the North-West (70.9%).

These differences produce an increase of the variability of the final weights and a decrease of the accuracy of estimated parameters.

6. Results

This paragraph illustrates the results of the analyses carried out on the distribution of main individual income sources singled out in the It-Silc survey: equivalent[6], employee, self-employment and pension (after retirement from employment) income. The said analyses have been carried out on the national territory and by geographical area according to the following hypothesis.

The first hypothesis analyzes the effects of attrition between the first and the second wave and compares 22.690 respondents and 3.216 attritors. The second hypothesis compares the attritors between the second and the third wave with

[6] The individual *equivalent income* is the total household income assigned to each of its members equivalized by the OECD modified scale.

20.012 respondents vs. 2.424 attritors. The third hypothesis, resulting from the combination of the two previous ones, analyzes the attrition between the first and the third wave and compares 20.012 respondents and 5.640 attritors.

Table 5 – *Evaluation of the effects of attrition between the first and the second wave by income and geographical area.*

Geographical area	Respondent				Attritors			
	Mean	Fi	Oi	Gi	Mean	Fi	Oi	Gi
North-West								
Equivalent income	18,169.73	0.5054	0.9999	0.2924	17,651.65	0.4669	0.9888	0.2692
Employee income	15,987.20	0.5033	1.0005	0.2860	15,664.61	0.4802	0.9908	0.2652
Self-employment income	16,133.32	0.5021	1.0031	0.4840	15,976.78	0.4876	0.9802	0.4679
Pension income	12,275.14	0.4977	1.0045	0.3223	12,002.56	0.5164	0.9688	0.2779
North-East								
Equivalent income	18,835.15	0.5000	0.9983	0.2978	18,411.57	0.5000	1.0153	0.2770
Employee income	15,249.10	0.5016	0.9986	0.3007	15,042.79	0.4896	1.0103	0.3049
Self-employment income	16,678.46	0.4952	1.0123	0.5066	16,786.68	0.5302	0.9165	0.3933
Pension income	11,488.78	0.4976	0.9982	0.3129	11,479.61	0.5205	1.0230	0.2943
Centre								
Equivalent income	17,154.20	0.5002	0.9982	0.2967	17,380.28	0.4990	1.0102	0.3083
Employee income	14,700.26	0.4996	0.9985	0.3045	14,794.55	0.5020	1.0084	0.3116
Self-employment income	14,226.38	0.4988	1.0087	0.4709	13,891.87	0.5056	0.9579	0.4297
Pension income	12,408.85	0.4978	1.0042	0.3383	12,809.95	0.5140	0.9726	0.3432
South								
Equivalent income	12,608.92	0.5007	0.9974	0.3072	12,843.27	0.4918	1.0309	0.3398
Employee income	12,683.24	0.5036	0.9974	0.3389	12,661.84	0.4603	1.0241	0.3687
Self-employment income	11,981.14	0.5017	1.0018	0.4832	11,918.59	0.4808	1.0064	0.5233
Pension income	10,409.51	0.4989	1.0007	0.3288	10,215.15	0.5137	0.9908	0.3626
Islands								
Equivalent income	12,665.20	0.4971	1.0044	0.3298	12,750.99	0.5259	0.9649	0.2988
Employee income	13,413.71	0.4994	1.0072	0.3616	13,482.82	0.5049	0.9470	0.3305
Self-employment income	12,661.71	0.5065	0.9866	0.4352	12,301.97	0.4434	1.1145	0.5303
Pension income	*no attritors for this sub-sample*							
Italy								
Equivalent income	16,463.76	0.4989	1.0017	0.3119	16,494.46	0.5080	0.9880	0.2996
Employee income	15,406.36	0.5014	1.0017	0.3040	15,243.50	0.4886	0.9876	0.3052
Self-employment income	14,887.75	0.4985	1.0055	0.4889	14,500.56	0.5093	0.9639	0.4470
Pension income	11,669.70	0.4971	1.0031	0.3281	12,023.74	0.5238	0.9734	0.3139

When comparing two sub-populations, the methodology described in section 4 undergoes a significant simplification. In order to verify the hypothesis that the two sub-populations come from the same population, or in other words, that there is complete overlap between them, the following conditions have to occur (Frick *et al.*, 2006, p.442-443):

i. $\overline{y}_{resp} = \overline{y}_{attr}$, same average income;

ii. $\overline{F}_{resp}(y) = \overline{F}_{attr}(y) = 0.5$, mean rank equals to 0.5;

iii. $G_{resp} = G_{attr}$, same Gini coefficient;

iv. $O_{resp} = O_{attr} = 1$, overlapping index equals to 1.

Table 6 – *Evaluation of the effects of attrition between the second and the third wave by income and geographical area.*

Geographical area	Respondent				Attritors			
	Mean	Fi	Oi	Gi	Mean	Fi	Oi	Gi
North-West								
Equivalent income	18,572.58	0.4981	0.9999	0.2924	18,747.05	0.5105	0.9989	0.3215
Employee income	16,851.38	0.5048	0.9968	0.2769	16,557.57	0.4745	1.0171	0.3102
Self-employment income	16,283.53	0.4990	1.0048	0.4818	16,535.53	0.5051	0.9786	0.4788
Pension income	12,097.54	0.5001	0.9949	0.3090	12,229.29	0.4994	1.0268	0.3514
North-East								
Equivalent income	19,277.84	0.5004	1.0026	0.3004	18,935.48	0.4966	0.9801	0.2843
Employee income	15,850.99	0.5020	1.0047	0.2974	15,332.98	0.4848	0.9583	0.2668
Self-employment income	17,568.23	0.5003	0.9991	0.5152	18,021.41	0.4982	1.0071	0.5357
Pension income	11,437.16	0.4983	1.0004	0.2985	11,635.57	0.5158	0.9973	0.2973
Centre								
Equivalent income	17,530.29	0.5025	0.9954	0.2911	17,173.44	0.4805	1.0350	0.3076
Employee income	15,469.71	0.5024	0.9998	0.2919	15,179.04	0.4817	1.0031	0.2969
Self-employment income	14,165.08	0.4978	1.0123	0.4734	13,936.29	0.5162	0.9071	0.3899
Pension income	12,702.85	0.5038	0.9960	0.3408	12,539.87	0.4661	1.0310	0.3302
South								
Equivalent income	12,880.00	0.4994	0.9956	0.3009	13,214.94	0.5081	1.0391	0.3232
Employee income	13,509.82	0.5028	0.9991	0.3337	12,901.04	0.4634	1.0086	0.3573
Self-employment income	11,971.85	0.4987	0.9938	0.4591	11,913.98	0.5154	1.0399	0.6489
Pension income	10,138.28	0.4981	1.0026	0.3072	10,297.75	0.5390	0.9612	0.2587
Islands								
Equivalent income	12,916.80	0.5008	0.9916	0.3300	13,739.76	0.4871	1.1196	0.4091
Employee income	14,086.05	0.5019	1.0004	0.3518	13,664.81	0.4696	0.9868	0.3348
Self-employment income	13,221.28	0.5037	0.9861	0.4740	14,010.70	0.4632	1.1531	0.5789
Pension income	10,583.31	0.4961	0.9899	0.3337	10,420.67	0.5835	1.0377	0.2851
Italy								
Equivalent income	16,745.74	0.4972	1.0006	0.3110	17,067.68	0.5234	0.9891	0.3315
Employee income	15,406.36	0.5014	1.0017	0.3040	15,243.50	0.4886	0.9876	0.3052
Self-employment income	15,094.98	0.4951	1.0188	0.4982	14,832.95	0.5126	0.9658	0.5130
Pension income	11,604.07	0.4988	0.9998	0.3186	11,954.37	0.5109	1.0012	0.3258

Any significant difference may be interpreted as an indication that the two sub-samples do not represent the same population. It may be better said that attritors' non-random behaviour can cause a bias in the analysis of income distribution.

Table 7 – *Evaluation of the effects of attrition between the first and the third wave by income and geographical area.*

Geographical area	Respondent				Attritors			
	Mean	Fi	Oi	Gi	Mean	Fi	Oi	Gi
North-West								
Equivalent income	18,063.40	0.5032	1.0025	0.2912	17,963.17	0.4914	0.9929	0.2852
Employee income	16,013.00	0.5063	0.9972	0.2816	15,889.76	0.4836	1.0050	0.2876
Self-employment income	15,852.88	0.4987	1.0090	0.4827	16,262.28	0.5035	0.9785	0.4795
Pension income	12,271.80	0.4962	1.0020	0.3227	12,151.00	0.5117	0.9963	0.2994
North-East								
Equivalent income	18,954.44	0.5007	1.0019	0.3019	18,508.99	0.4975	0.9926	0.2709
Employee income	15,344.13	0.5045	0.9965	0.3006	14,818.04	0.4852	1.0108	0.3033
Self-employment income	16,884.64	0.4956	1.0184	0.5125	16,537.06	0.5138	0.9387	0.4211
Pension income	11,461.74	0.4946	0.9946	0.3146	11,600.94	0.5234	1.0016	0.2945
Centre								
Equivalent income	17,213.96	0.5023	0.9970	0.2959	17,112.32	0.4930	1.0095	0.3063
Employee income	14,701.10	0.5025	0.9949	0.2997	14,755.93	0.4928	1.0148	0.3229
Self-employment income	14,314.45	0.4977	1.0182	0.4789	14,700.21	0.5063	0.9478	0.4195
Pension income	12,497.96	0.5003	1.0016	0.3404	12,460.68	0.4990	0.9951	0.3342
South								
Equivalent income	12,548.35	0.4993	0.9953	0.3050	13,020.27	0.5042	1.0270	0.3378
Employee income	12,698.68	0.5052	0.9979	0.3373	12,021.77	0.4703	1.0104	0.3644
Self-employment income	11,897.72	0.4989	1.0040	0.4847	11,936.79	0.5059	0.9830	0.4991
Pension income	10,397.02	0.4969	1.0027	0.3303	10,999.12	0.5231	0.9777	0.3389
Islands								
Equivalent income	12,628.20	0.4957	0.9980	0.3269	12,919.76	0.5233	1.0171	0.3235
Employee income	13,524.38	0.5025	1.0054	0.3610	12,922.74	0.4880	0.9721	0.3426
Self-employment income	12,639.82	0.5048	0.9895	0.4405	12,403.42	0.4802	1.0496	0.4642
Pension income	10,684.04	0.4935	1.0117	0.3371	10,455.45	0.5484	0.9312	0.2555
Italy								
Equivalent income	16,389.32	0.4957	1.0047	0.3128	16,749.10	0.5156	0.9822	0.3013
Employee income	14,695.65	0.5021	0.9994	0.3093	14,557.87	0.4929	1.0021	0.3144
Self-employment income	14,843.96	0.4961	1.0119	0.4931	14,806.90	0.5124	0.9616	0.4514
Pension income	11,656.71	0.4949	1.0042	0.3299	11,922.03	0.5211	0.9826	0.3127

A first analysis of results obtained shows a certain homogeneity compared to the specified three work hypotheses; it emerges that the sample of respondents and

that of attritors come from the same population or, in other words, are representative of the same population.

Another element common to all the three work hypotheses is the substantial difference among income distributions that, although predictable, is unlikely to be observed both in the sample of respondents and in that of attritors.

7. Conclusion and remarks

The ANOGI is used on It-Silc data in order to analyze non-response behaviours and in the evaluation of the effects of attrition on the core survey variable (i.e. income). In particular, it proves the efficacy and simplicity of use of the ANOGI within the variability study of sub-populations.

Unlike several studies on attrition mainly aimed at determining the response probability in function of individuals' characteristics and of the context in which the survey is conducted, this study introduces a new perspective by the direct evaluation of the attrition effect on studied variables.

The results of this work represent a first step towards the realization of a larger project mainly aimed at studying methods for the analysis of income, living conditions and poverty.

To briefly sum up the results obtained, it can be said that the panel drop-out, analyzed separately between the first and the second wave, and between the second and the third wave and on the whole panel, results in an increase of sampling error.

The supposed bias due to sample self-selection produces only negligible effects on all types of income and geographical areas.

Acknowledgements

The present work has been realized within the PRIN-2007 project: "*Inequality, poverty and social exclusion: analysis and coherence of information sources, new measures and interpretative methods*".

References

Ceccarelli C., Cutillo A. 2007. *Il trattamento della mancata risposta totale nell'indagine Eu-Silc: una valutazione tramite una misura del cambiamento,* Congiuntura**,** 1° Trimestre 2007, CREF, Udine, 91-112.

De Vergottini M. 1950. *Sugli indici di concentrazione,* Statistica**,** 10, 445-454.

Deville J. C., Särndal C. E. 1992. *Calibration Estimator in Survey Sampling,* Journal of the American Statistical Association, 87, 376-382.

European Parliament and Council 2003. *Regulation (EC) No 1177/2003 of the European Parliament and of the Council of 16 June 2003 concerning Community statistics on income and living conditions (Eu-Silc)*, Official Journal of the European Union.

Frick, R. J., Goebel, J., Schechtman, E., Wagnar, G.G., and Yitzhaki, S. 2006. *Using Analysis of Gini (ANOGI) for Detecting Whether Two Sub-Samples Represent the Same Universe*. *The German Socio-Economic Panel Study (SOEP) Experience*, Sociological Methods and Research, 34, 427-468.

Gini, C. 1912. *Variabilità e mutabilità: contributo allo studio delle distribuzioni e delle relazioni statistiche*. In: Studi Economico-giuridici della Regia Facoltà di Giurisprudenza, 3(2), Bologna: Cuppini.

Gini, C. 1914. Sulla *misura della concentrazione e della variabilità dei caratteri,* Atti del Reale Istituto Veneto di Scienze, Lettere ed Arti, 73, 1203–1248. (English translation in Metron, 2005, 63, 3–38).

Giorgi, G. M. 1990. *Bibliographic portrait of the Gini concentration ratio*, Metron, 48, 183–221.

Giorgi, G. M. 1992. *Il rapporto di concentrazione di Gini: Genesi, evoluzione ed una bibliografia commentata*, Siena: Libreria Editrice Ticci.

Giorgi, G. M. 1993. *A fresh look at the topical interest of the Gini concentration ratio,* Metron, 51, 83–98.

Giorgi, G. M. 1999. *Income Inequality Measurement: The Statistical Approach*. In J. Silber ed., Handbook on Income Inequality Measurement, Boston: Kluwer Academic Publishers, 245-260.

Giorgi, G. M. 2005. *Gini's Scientific Work: An Evergreen*, Metron, 63, 493-503.

Istat 2008. *L'indagine europea sui redditi e le condizioni di vita delle famiglie (Eu-Silc),* Collana Metodi e norme, n.37.

Kass, G.V. 1980. *An explanatory technique for investigating large quantities of categorical data*, Applied Statistics, 29, 119-127.

Lerman, R. I. and Yitzhaki, S. 1984. *A Note on the Calculation and Interpretation of the Gini Index*, Economics Letters, 15, 163–68.

Osier G., Museux J.M., Seoane P., Verma V. 2006. *Cross-sectional and longitudinal weighting for EU-SILC rotational design,* contributed paper on Methodology of Longitudinal Survey (MOLS 2006), University of Essex.

Piesch, W. 1975. *Statistische Konzentrationsmasse,* Tübingen: J.B.C. Mohr (Paul Siebeck).

Pyatt, G., 1976. *On the interpretation and disaggregation of Gini coefficient,* Economic Journal, 86, 243–255.

Rendtel U. 2002. *Attrition in Household Panels: a survey*, Chintex working paper, 4.

Stuart, A. 1954. *The correlation between variate-values and ranks in samples from a continuous distribution,* British Journal of Psychology, 7, 37–44.

Yitzhaki, S. 1988. *On Stratification and Inequality in Israel,* Bank of Israel Economic Review, 63, 36-51.

Yitzhaki, S. 1994. *Economic Distance and Overlapping of Distributions,* Journal of Econometrics, 61, 147–159.

Yitzhaki, S. and Lerman, R. 1991. *Income Stratification and Income Inequality*, Review of Income and Wealth, 37, 313–329.

SUMMARY

The Italian National Institute of Statistics (Istat) has set up a survey on income and living conditions (It-Silc), mainly composed of a cross-sectional and a longitudinal component.

This paper aims at proving, by a decomposition of Gini concentration index, also known as Analysis of Gini (ANOGI), whether attrition introduces an element of bias in the analysis of income distribution.

Compared to other studies in the literature, it introduces a new perspective by the direct evaluation of the attrition effect on studied variables.

Claudio CECCARELLI, Senior Researcher of Statistics, Italian National Institute of Statistics, Division for Survey on Living Conditions and Quality Life, clceccar@istat.it

Giovanni Maria GIORGI, Full Professor of Statistics, University of Rome "La Sapienza", Department of Statistics, Probability and Applied Statistics, giovanni.giorgi@uniroma1.it

Rivista Italiana di Economia Demografia e Statistica Volume LXIII nn. 1-2 – Gennaio-Giugno 2009

POVERTY AS LACK OF OPPORTUNITY: A COMPARISON BETWEEN JOHN ROEMER AND AMARTYA SEN

Enrica Chiappero Martinetti

1. Introduction

Economic inequality has traditionally been conceived as inequality of outcomes, measuring income, consumption or wealth differences among individuals or families. However, a considerable controversy still surrounds the main sources of economic inequalities; the extent to which they can affect other outcomes; whether or not these disparities matter.

In the last decade, this debate has increasingly shifted from inequality of outcome to equality of opportunity. This approach paid greater attention to social and economic circumstances which hinder individuals from competing at the same starting level, and to the real opportunities people have to pursue their own life plans and objectives.

Among others, Amartya Sen and John Roemer are considered the authors who made a significant contribution to the debate, even though they started from different premises, and reached rather different conclusions. According to Sen, individual overall advantage should not be assessed in terms of resources or utility but in terms of the person's capability to do things he or she has reason to value[1]. Various set of contingencies, including personal characteristics, social and economic circumstances, can affect the real opportunities people have, generating variations in the conversion process of income and economic resources into achievements. These sources of variations are crucial in understanding poverty and inequality and should be taken into account when designing public policy addressed to it.

Similarly, John Roemer triggered an engaging debate on this issue in his book *Equality of Opportunity* (1998), remarking that individuals should not be held responsible for those circumstances which are clearly beyond their control (such as

[1] "The capability approach focuses on human life, and not just on some detached objects of convenience, such as incomes or commodities that a person may possess, which are often taken, especially in economic analysis, to be the main criteria of human success. Indeed it proposes a serious departure from concentrating on the *means* of living to the *actual opportunities* of living" (Sen 2009, p. 233).

gender or race). Inequality of outcomes should reflect only differences due to individual choices, like effort, or different talent. Unequal opportunities caused by circumstances at birth, broadly recognized as unfair by people, should be reduced or eliminated.

In this paper, I will compare and contrast the two approaches to the problem, pointing out analogies and differences between Roemer and Sen's views. Section two will be devoted to briefly discuss Roemer's equality of opportunity, while in section three I will analyse Sen's equality of capabilities. In section four, I will compare and contrast the two theories. Section five will conclude the paper.

2. Equality of opportunity

Drawing on Arneson's earlier formulation (Arneson 1989), Roemer (1998) further developed and formalised the idea of equal opportunity in the quest for well-being. According to him, success in life should not depend on birth circumstances. Rather, it should centre exclusively on people's autonomous choices, effort and personal skills. Features such as gender, race or social background, which are beyond the individual's control, cause a set of morally unacceptable inequalities. With reference to outcome (which he defines as advantage), Roemer points out how personal commitment and individual responsibility play a major role in performing the actions required to achieve well-being. In a society where equality of opportunity is the rule, relevant outcomes (including income, wage-earnings capacity and life expectancy) are redistributed regardless of the circumstances already mentioned. Therefore, effort is the residual which explains outcome differences. If this is not the case, as the existence of profound disparities observed in most societies testify, an equal opportunity policy should be oriented to level the playing field among people.[2]

Equal opportunity literature occupies a broad field of philosophical and economic research, including contributions by Cohen (1989), Fleurbaey (2005), Fleurbaey e Maniquet (2007, 2009), Sen (1985) and Roemer (1998, 2002, 2009) who gave a substantial contribution to translate these philosophical ideas into an

[2] Roemer recently suggested (2009) the possibility to extend the equality-of-opportunity idea to intergenerational and global equality. In the first case, the need for equality is based on the assumption that individual welfare should not be affected by the date at which the individual is born. Thus, a sustainable path of development should allow maximizing the welfare level that can be enjoyed by *every* generation. In the second case, according to cosmopolitanism (Pogge 2002), inequalities among individuals of different nationalities shall be considered morally arbitrary because birthplace is also a circumstance beyond individual responsibility.

economic algorithm for measuring equality of opportunity through statistical analysis.

Roemer specifically assumes that the outcome observed (*u*) is determined by a set of circumstances (*c*), affecting a certain group of people or "type" *t*; the effort (*e*) which comprises the actions people take and for which they are responsible; the set of policies (*p*) used to equalize opportunities, so that it is possible to write *u (c, e, p)*.

For instance, if wage-earning capacity is the objective (e.g. "the condition of individuals whose acquisition we desire to equalize opportunities for", Roemer et al. 2003, p. 541), such a capacity certainly depends on the individual's effort as well as on circumstances like the parental socioeconomic status – loosely defined according to one's parents' education level – which may play a relevant role in improving this capacity. A policy aimed to provide equal opportunities could involve investments in educational resources, so that future earning capacity of different individuals might be equalized. Similarly, if life expectancy is the objective, effort could be defined as the individual's commitment to maintaining a healthy lifestyle in accordance with the *t* group specific features (if gender is the relevant circumstance, women might have healthier life-styles than men). In this case, a type-specific policy aiming to equalize life expectancy could entail health care campaigns specifically directed towards the more disadvantaged group, in order to raise their life expectancy to healthy life-style levels for each type. What matters here is that the distribution of effort among individuals of type *t* is a characteristic of the type, not of any individual.

When putting the theory into operation, the most difficult problem concerns the possibility to estimate individual effort which is the result of a process of individual preference maximization**,** and therefore not easily observable. As Hild and Voorhoeve (2004) point out, Roemer's definition and measurement of effort varied through the years. As already said, in his 1998 book Roemer assumes that inequalities are due to different choices in terms of the relative effort, given the circumstances individuals experience. Therefore, "individuals should be held accountable for their degrees of *effort* but not their levels of *effort* […] I say it is morally wrong to hold a person accountable for not doing something that it would have been unreasonable for a person in his circumstances to have done" (Roemer, 1998, p.18).

This conception has been recently modified (see Roemer 2009 and Roemer et al. 2003). According to the new version, individuals relative effort not only is determined by choice-related variables, but also by all factors which determine individuals' rank in their type's outcome distribution, including personal skills and talents, as well as good luck. What matters to Roemer is finding an "inter-type-comparable effort measure" (Roemer et al. 2003, p. 543)

The underlying idea is that individuals of different types who occupy the same rank in their outcomes distribution functions engage with their duties at the same level and thus, they behave in an equally responsible manner even if their absolute levels of effort are different. It would not be fair, for instance, to compare school performances by two pupils who grew up in very different living contexts. Rather, their levels of effort should be compared with others' experiencing the same set of family circumstances. If the two children had the same propensity to apply themselves, they should receive equal rewards or advantages, while if one of them is a member of a disadvantaged group this will affect his effort level. In other words, the distribution of effort becomes itself a characteristics of the type, not of the children. Roemer (2002) proposes as a measure of the relevant degree of effort the quantile of the effort distribution for the type an individual belongs to.

Two general approaches have been suggested for equalizing the advantage of all individuals across types. Roemer suggests a mean-of-mins or sum-of-mins rule, which states that the selected policy p within the possible social options $p \in \Phi$ should maximize the minimum level of outcome (U) achieved by individuals (i) in the same circumstances (I = 1…T), given their diverse effort levels (j=1,…,N)

$$max_{p \in \Phi} \sum_{j \in N} min_{i \in T} U^{j}_{i}(x)$$

A different approach, defined by Dirk van de Gaer as min-of-means or min-of-sums rule (Van de Gaer, 1993; see also Kranick, 1996; Ok and Kranick, 1998, Bossert et al. 1999), focuses on the set of outcome available to the members of each type (their opportunity set). It takes the average outcome of each type and requires equating the set of opportunities, maximizing the most disadvantaged one:

$$max_{p \in \Phi} min_{i \in M} \sum_{j \in N} U^{j}_{i}(x)$$

The two rules are very similar and Roemer (2002, p. 459) considers them as two possible options among the several available. In fact, there is an apparent relation between the two. The opportunity set held by an individual in specific circumstances is associated to a vector of possible outcomes. Their achievement varies according to the diverse effort levels the individual engages with. So, if the same group is the most disadvantaged at each effort level, the two rules will lead to the same policy p.

However, Ooghe, Schokkaert and Van de Gaer (2007) point out how the differences between the rules go beyond formal elements. The "compensating outcome" approach suggested by Roemer argues that equalizing opportunities

entails equalizing the individual outcome. In fact, this latter mirrors differences among types and circumstances which are beyond individual control – and responsibility. In this interpretation, equalizing opportunities means also accepting inequalities in the individual outcome, if they derive from a different degree of effort. In other words, given equality of opportunity, the same degree of effort should produce equal individual outcomes. When this does not occur, individuals should be compensated, only if circumstances differ. What matters here is therefore equality *among* groups, not *within* the same group.

On the other hand, in Van de Gaer's "compensating sets" approach compensation refers to the different *opportunity sets* available to individuals and therefore aims to equalize the value of these sets. However, as Ooghe, Schokkaert and Van de Gaer show, they have a very different ethical inspiration and can lead to different results in terms of policy prescriptions (see also Moreno-Ternero, 2007).

2.1 *Some open issues about equality of opportunity*

The equality of opportunity is currently one of the most supported conception of distributive justice and has greatly influenced the debate in egalitarian political philosophy. The literature about the measurement of inequality of opportunity, both at theoretical and empirical level, flourished during the recent years and produced very interesting results (see among others, Ruiz Castillo, 2003; Peragine, 2004, 2005; Savaglio e Vannucci, 2007; Lefranc, Pistolesi and Trannoy, 2006). Beside these remarkable results, there are also some unresolved questions, as it reasonable to expect. I would like to raise four of these issues in the current section.

The first issue concerns a possible overlapping between opportunities and circumstances. If equality of opportunity was an achieved and well established fact, this would entail an equal set of circumstances. Therefore, according to Roemer**,** outcome inequalities would be solely due to different levels of effort and thus being considered morally acceptable and fair. On the contrary, unrelenting and broad inequalities affecting modern societies seem to demonstrate that unequal opportunities deriving from a diverse set of circumstances persist. These latter, in turn, will originate different outcome even in case of equal levels of effort. In such cases, if we can observe circumstances and outcomes but we cannot measure effort, how can we distinguish between equality and inequality? By and large, opportunities and circumstances tend to overlap the more we move from the conceptual to the empirical sphere.

This overlapping, which can generate major consequences on estimation as we shall consider later, is not neutral also at conceptual level. In the first place, other

circumstances being equal, every possible outcome level is theoretically compatible with an equality-of-opportunity principle. In other words, inequality of opportunity is consistent with every possible unequal outcome. Moreover, in most cases circumstances like gender, race or family background cannot be changed. Therefore, how can we guarantee equality of opportunity in these cases? What kind of interventions can be implemented to restore it? Are there any policies capable of accounting for different circumstances and compensating for them? Finally, is it always possible to justify such policies?

Consider the examples represented in figure 1. Suppose that there are only two types, A and B, the former experiencing better circumstances. As already pointed out, we can observe their outcome but not their effort level. Let us also assume, as in the first row, that the outcomes observed are equal. Since circumstances are different, we can also infer that the effort levels in each group are different and, more specifically, the more advantaged certainly put less effort into his action. If we take equal opportunity seriously, we should reward the B individual for his higher effort which allowed him to achieve a higher outcome despite his adverse circumstances. Or, alternatively, should we penalize the more advantaged person for his lower level of effort?

Figure 1 – *Equality of Opportunity. A Simple Example.*

Circumstances	**Effort**	**Outcome**	**Policy Implications**
$C_A > C_B$	$e_A < e_B$	$u_A = u_B$	Shall we compensate B for the greater effort or penalize A for most favorable circumstances?
$C_A > C_B$	$e_A > e_B$ $e_B \geq e_A$	$u_A > u_B$	Unequal outcome is legitimate. But what owes to *c* and what to *e*? Shall we compensate B or penalize A?
$C_A > C_B$	$e_B \geq e_A$	$u_A < u_B$	Shall we give an additional reward to B? How much additional effort is required for compensating the less favorable circumstances?

Let us consider the second row where the more advantaged person A gets the better outcome. At a glance, this may seem legitimate because of his higher effort level. However, we can observe that A also experiences better circumstances. How then can we distinguish between higher effort and better circumstances when accounting for a better performance? Alternatively, consider the case where A put less effort into his action. We should again compensate B for his higher effort, even though this latter is not sufficient to make up for adverse circumstances and to achieve a better outcome. The third case, described in the last row, has similar implications in terms of policy.

Whether the outcome is earning capacities, school performances or life expectancy, we always end up with paradoxical conclusions, at least unlikely to sustain or justify. Of course, we can award scholarships to a pupil with outstanding performances and a bad family background. But, is it right to penalize a pupil with the same outstanding performances and a good family background, if all this does not depend on higher effort? What if his lower effort were due to his greater talent?

Two reasons account for this problem, I think. First, the overlap between opportunities (non-observable) and effort measurement in the relation between circumstances and outcomes. Second, the division into groups or types, namely the criteria to be used. If they are defined according to outcome differences, how can we be sure B is actually more disadvantaged than A? What if such disadvantage depended on a lower effort level? Consider women's consistently low wage levels in the labor market. Among the diverse reasons provided to explain it in the literature (for instance, gender discrimination), sometimes the low effort put into work is also used to account for it. How can we measure such effort, if the type taken into consideration is itself defined according to gender-based prejudices affecting the type's average outcomes? Labor economics literature and feminist economics describes gender discrimination on labor market as a consequence of stereotypes and lack of signals which allow to identify the actual level of productivity. Thus explained, they implicitly reaffirm women inequalities in terms of opportunity as well as of outcome, paying no attention to their effort. How is it possible to intervene in these cases?

A second issue concerns where the "starting gate" should be placed, i.e. how to define the circumstances from which to level the playing field, in order to achieve equality of opportunity. What sort of differences should not be levelled off? This problem links to a broader issue, namely what circumstances must be preserved in order to achieve equality of opportunity. Who selects them? Which among them are actually beyond individual control? A wide range of answers is possible. The minimal version includes in the list only gender, race or religion. Broader lists also include social and family background. Finally, the maximal version takes into account every difference linked to birth and chance, considering choices and preferences the only morally legitimate sources of outcome inequalities.[3]

In short, distinguishing between circumstances and actual effort is not an empirical question but a morally relevant action. In addition to this, such distinction is often based on pre-existing unequal opportunities. Is talent a morally justified basis for accepting outcome differences? Does it originate from a natural lottery? Is it a genetic factor? Does it depend on environmental circumstances?

---

[3] Hild and Voorhoeve (2004) define the three inclusion levels as *right-liberal, left-liberal* and *socialist catalogue*. See also Cohen (2001) and Kolm (2001).

A third important issue regarding effort has been pointed out by Fleurbaey (1998). According to him, Roemer's definition of relative effort can allow to separate morally relevant factors (i.e. effort) from morally irrelevant ones (i.e. circumstances such as race or gender) if and only if the former are not statistically correlated to the latter. Looking at the life expectancy example, one may argue that a different disposition towards a healthy lifestyle is significantly correlated with the information available to each group on the risks associated to their actual lifestyle. This refers not only to the amount of information, but also to its appropriateness.

This point is of paramount importance when we put the equality of opportunity theory into operation. For instance, a recurrent criterion for partitioning population into groups or types is parents' education as a proxy for the individual's family and social background, thus defining the circumstances. The implicit assumption is that the residual (that is, the effort) accounts for different outcome levels. One may notice that genetic differences, talents and effort are correlated with the parents' education level – as they usually are (Saunders 1996). A problem of endogeneity immediately arise when distinguishing opportunity, effort and circumstances from outcomes is difficult. Such a problem is more relevant with reference to adult people while is relatively less serious in case of children for whom is reasonable to assume that their access to opportunities (not to outcome) is independent to the effort.

Finally, Roemer's version of equality of opportunity assumes that what has to be levelled off is an average outcome - in terms of resources, education levels or life expectancy – for each type. However, as pointed out earlier, there is a trade-off between number of cells, and thus accuracy of type's definition, and sample size and therefore, the sampling variance within each cell which depends on the number of observations. As Ferreira and Gignoux notice (2008), this may lead to an over-estimation of inequality of opportunity. To avoid this, the samples used should be representative enough to allow an accurate types' specification. Alternatively, the authors suggest integrating non-parametric analyses with parametric estimations. This allows them to specify a lower-bound estimate of inequality of opportunity, rather than a single index.

3. Equality of capability

In this section I will briefly consider Amartya Sen's equality of capability before considering in section 4 some interesting analogies with Roemer's equality of opportunity, as well as their significant differences.

Since the first appearance of *Commodities and Capabilities* in 1985, many scholars from a broad range of disciplines have been analysing and expanding

Sen's capability approach, including the distinguished contributions of Martha Nussbaum. Nowadays, the existing literature on the subject is immense, involving very diverse fields of investigation. Even if it took many years to account for concepts like *functionings* or *capabilities*, whose groundbreaking value often remained ambiguous to many, these expressions are now part of the vocabulary of economists and philosophers. Therefore, I will just make a brief overview of the basic concepts of this approach.

According to Sen, *"a functioning is an achievement, whereas a capability is the ability to achieve. Functionings are, in a sense, more directly related to living conditions since they are different aspects of living conditions. Capabilities, in contrast, are notions of freedom in the positive sense: what real opportunities you have regarding the life you may lead*" (Sen 1987, p. 36). Sen suggests replacing access to resources with a broader notion of well-being, namely the one including "what an individual can do and can be." In this sense, individual well-being does not depend exclusively on the resources available to people. Rather, the definition includes their ability to transform these resources into valuable achievements. A *capability set* consists of the possible valuable achievements, whereas *functionings* describe those actually achieved.

The capability approach does not simply contrast with the more traditional income-based approach, but encompasses it. Material resources such as income are of paramount importance in determining well-being, but they are not a sufficient metric. According to Sen, a complete definition of well-being also includes achievements whose value cannot be determined by money. Sen lists among them an adequate education; a good health and nourishment level; bodily integrity against any from of violent assault; control over one's material, political and cultural environment.

Another important distinction in Sen's argument lies between the notion of *standard of living*, as opposed to *well-being*. The first refers to what has a direct impact on individuals personal lives. The second takes into account all those aspects contributing to individual well-being, even though they are not directly linked to it. In this sense, adequate nourishment falls into the first category, whereas working actively for famine relief falls into the second one. Linked to this distinction is the notion of *agency*. Sen defines it as people's actual possibility and capability to pursue tasks they have reason to value, whether or not they may affect their standard of living.

Finally, the concept of freedom is pivotal when looking at Sen's theory. In his approach, freedom is above all an individual's actual ability to choose freely what to do, which tasks to pursue, in short what kind of life is worth living. In this sense, freedom is substantive rather than instrumental, thus giving body and value to the idea of development and well-being.

It would be appropriate to look at people's standards of living as a multi-dimensional phenomenon, even when our attention is directed to very specific aspects of it – for instance, to a single agency possibility, or to a restricted set of functionings. In addition to the complex definition of well-being already explained, Sen's most outstanding contribution to the problem is to offer a unifying theoretical framework for approaching welfare economics, so that a wide range of issues traditionally separated can be brought together.

Of course, there are some objective difficulties when trying to operationalize this complex and challenging theoretical framework, namely in formulating algorithms, measures and policies based on it and setting up tools to assess and control their efficacy. It is also true that in the last ten years a growing body of empirical studies referring to the capability approach have developed and interesting methodologies and tools suggested.

2.2 *Some open issues about equality of capabilities*

Sen's capability approach has been given a huge attention and a broad consensus. Most scholars appreciate its innovative nature, offering an original and broad perspective for conceptualizing and understanding poverty, inequality and development; its attempt to shift the attention from the means of living to the actual opportunities of living; its ability to bring together aspects such as freedom or agency that were hitherto excluded from (or inadequately formulated in) traditional approaches to welfare economics; its capacity to represent a genuine alternative paradigm for measuring well-being, compared to standard income-based methods. However, other authors consider these innovations as weak points of this theoretical framework, for several reasons.

Critics argue that Sen's approach is too complex, on one hand, and on the other not enough defined, leaving unresolved specific normative questions. Sen himself acknowledges that the capability approach is not aimed to be a normative theory on social justice, even though it can serve as an important constituent for it (Sen, 1995, 2004; see also Robeyns, 2005). This kind of theories would generally require i) an aggregative principle from individual well-being to that of society as a whole; ii) a procedure to formulate the method, which the capability approach does not specify[4]. Despite its underspecified nature, the capability approach still maintain its relevance for assessing and comparing social disparities. In his new book *The Idea of Justice* (Sen, 2009), Sen has recently argued that, above all, a theory of justice

[4] Martha Nussbaum's contributions (2000, 2003) can be seen as a step towards a theory of justice based on the capability approach by specifying a list of human capabilities.

should aim to reduce grave iniquities in the world, rather than wonder about what a hypothetical perfectly just society could or should be like. While it is almost impossible to agree on what perfect justice is, it is feasible to acknowledge actual injustices, like slavery, women's subjugation, famine and poverty. In these cases, according to Sen, we must intervene and the capability approach can offer an important informational focus for assessing freedom and justice.

There are, however, other important open issues to be considered. One could be which capabilities should be included, and whether there should be a scale of priorities. Sen considers the capability approach as a flexible pattern, open to diverse conceptions of good, justice and advantage. Therefore, it is not appropriate to make an *apriori* list. Rather, one should be built according to any given society's culture and values, thus varying on the basis of time and space.

On the contrary, Nussbaum strongly argues in favour of a universal list of capabilities. She identifies ten central human capabilities, each of which is at same level and cannot be replaced or compensated by anything else. This list should be a founding principle in the constitution of each Country, which should promote and defend it for every individual in every society. However, there is still controversy among scholars over the procedure to be used in order to define such a list.[5]

Despite the difficulties, the recent experience conducted by the Equality and Human Rights Commission in the UK testifies the concrete possibility to implement the capability approach and formulate a policy framework substantially based on it[6]. In particular, this Commission defined a set of ten central and valuable freedoms drawing on universal human rights framework, as well as on deliberative consultation with the general public and individuals and groups at high risk of discrimination and disadvantage. The Commission also suggested that these dimensions of equality should be jointly considered and comparisons made across the whole set of dimensions.

A controversy still remains with reference to poverty issues broadly considered.

---

[5] Drawing on the notion of basic needs, Alkire (2002) suggests to start with a sufficiently complete list of capabilities able to generate the broadest consensus. Robeyns (2003) proposes a procedure and a set of criteria to be applied for minimizing biases and reach a consensus. Other authors remark that we should focus on a small set of universally recognised dimensions, such as health, education and housing conditions which are also traditionally considered in most sample surveys on quality of life. One may object that most empirical studies in this field make use of data which are not gathered for measuring functionings or capabilities, and no normative justification is provided with reference to their relevance. See Alkire (2002), Gasper (2004), Robeyns (2005) and Giovanola (2007) and Magni (2006) on ethical and philosophical aspects related to the list.

[6] See Equalities Review, 2007a e 2007b; Equality and Human Rights Commission, 2008; Burchardt e Vizard, 2007.

In fact, it is not clear whether within a given list of capabilities (or independently from its existence) a special attention should be paid to a set of basic capabilities (which can include among the others an adequate nourishment and health standards, access to education, social and political life) which cannot be disregarded, no matter what the degree of inequality is in other dimensions.[7]

Here I want to point out that arguing over the number of dimensions to be taken into account in the analyses (alternatively linking to poverty, inequality or well-being) is a problem shared by every multidimensional approach, including the equal opportunity approach discussed in the previous session. It seems, however, that only the capability approach literature has developed a serious debate on it.

A second contentious point refers to the concrete possibilities to operationalize the capability approach due to a lack of a specific metric for valuing capabilities or functionings, as well as to its rather demanding structure, because of the quantity and quality of statistical information required. In fact, as already mentioned, unlike Roemer Sen does not provide a mathematical formula or an algorithm for measuring capabilities or functionings. This choice is consistent with the original aims of this approach which want to be a broad, flexible "way of thinking" adaptable to different issues in all possible contexts. From this point of view, the main problem does not seem to suggest a single formula, but rather to test different strategies and techniques suitable to put such theoretical framework into operation, preserving its original richness.

Sen neither ignores nor underestimates these difficulties. The capability framework he suggests is intrinsically pluralist not only with reference to the number of dimensions to be measured, but also in terms of heterogeneity of the agents as well as of the multiplicity of contexts under analysis. On one hand, the standard poverty measurement approach (based on an income poverty threshold used as a benchmark to compare and contrast individual life conditions) is totally inadequate. As Sen argues, in order to take human diversity and heterogeneity of contexts seriously, even if we want to preserve a money metric (something which represents nevertheless a second-best option) we should not refer to a minimum income threshold applicable to all. Rather, we should be able to identify adequate thresholds which allows people to achieve a minimum set of functionings, according to their characteristics and to the specific circumstances they live. On the other hand, if we move from a unique metric to a genuine multidimensional view

[7] It is important to notice that Nussbaum and Sen do not refer to "basic capabilities" in the same way. Sen defines them as a set of *beings* e *doings* which make human life dignified, such as the essential ones just mentioned. Nussbaum describes them as "a set of innate abilities which constitute a basis for developing advanced skills" (Nussbaum, 2000). Unfortunately, the term "capability" is ambiguous in itself and sometimes Sen seems to use it in its literal meaning, i.e. an individual ability to do. See Sen (1993).

some additional problems of arbitrariness arise, including the fact that there is no unique broadly acknowledged method of aggregation across dimensions or a pre-defined set of weights to be assigned to each of them. Alternatively, we could simply rank some vectors vis-à-vis others, instead of collapsing multidimensionality into a synthetic index of poverty or well-being. This allows for partial instead of complete orderings, but as Sen remarks it makes more sense to accept less ambitious partial ranks than to insist on arbitrarily complete orders (Sen, 1985).

Sen's approach is also considered very demanding because of the range of information required. This is true, although not very different from other multidimensional approaches to poverty. However, the growing availability of datasets stored and spread by International Agencies and Statistical Offices as well as some interesting examples of ad hoc analysis conducted in some countries are gradually coping with the lack of data, providing a broad range of good quality data.

Nowadays, Sen's approach has been used in an increasing number of empirical studies (see Chiappero-Martinetti e Roche, 2009 for a recent review), whose innovative methods and analytical rigour implicitly confirmed its validity, showing that measuring capabilities and functioning remains a challenging but feasible exercise.[8]

Moreover, it is important to remark that most of these problems do not pertain exclusively to the capability approach but affect every multidimensional approach to poverty and well-being analyses.[9]

A third issue concerns paternalism, namely the approach's tendency to allow

[8] See Eurostat and OCSE Internet websites for EU countries. With reference to developing countries, there is now a growing availability of household surveys: e.g. see the World Bank LSMS Project - Living Standard Measurement Study (www.worldbank.org/lsms). The HDCA website provides an extensive review of both empirical analyses and methodological tools used for operationalizing the capability approach. See: www.capabilityapproach.org

[9] See Atkinson (2003), Bourguignon e Chakravarty (2003). In Sen's words (1992, p. 108-109): "If we concentrate on certain basic general functionings and corresponding capabilities, there may be much more agreement on their importance than there would be if we concentrate on particular commodity bundles and particular ways of achieving those functionings. For example, there is likely to be more intercultural – and also interpersonal – agreement on the importance of having the capability to avoid acute hunger or severe undernourishment, than on the significance of having an adequate supply of particular food items (e.g. some specific type of meat or fish or grains or pulses) to serve those functionings […]. This is, in fact, one reason why poverty is better seen in terms of capability failure than in terms of the failure to meet the 'basic needs' of specified commodities."

public policies to interfere with individual choices. When discussing the link between personal and social responsibility, Sen underlines how the former should not replace the latter, because responsibility presupposes freedom. Social and individual responsibility can and must complement each other. As he writes (1999, p. 284): "The adult who lacks the means of having medical treatment for an ailment from which she suffers is not only prey to preventable morbidity and possibly escapable mortality, but may also be denied the freedom to do various things – for herself and for others – that she may wish to do as a responsible human being." Promoting society's interventions towards an increasing individual freedom therefore enhances individual responsibility. According to Sen, there is a strong social obligation to guarantee those circumstances allowing human capabilities to be fully expressed. At the same time, this being an opportunity-based approach, individual choices and responsibilities maintain a pivotal role.[10]

In this respect, Nussbaum's version is perhaps weaker because of the predetermined list of capabilities, which in turn presupposes an *apriori* conception of the good. Nussbaum argues that the list's contents are far too general and minimal to be paternalistic. Rather, they shall be used to build a basic consensus area among different conceptions of the good. Moreover, referring to capabilities rather than to functionings leaves more room for individual choices and responsibilities.

4. Equality of opportunity vs. equality of capability

Equality of opportunity bears a strong similarity to equality of capability.[11] Both the approaches distance themselves from equality of resources and equality of outcomes and take an intermediate position in the abstract process of generating well-being. This latter can be expressed as a sort of production function which transforms inputs (resources, public and private goods) into outputs (outcomes,

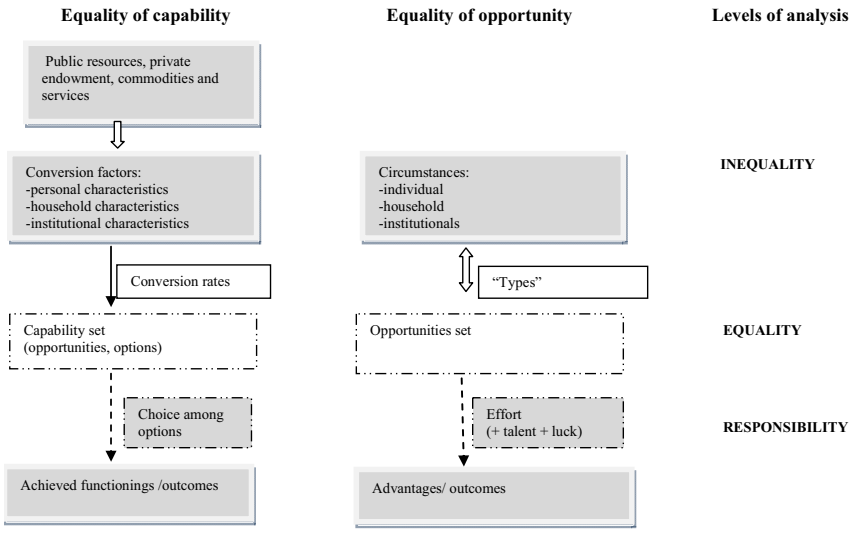
[10] "The denial of opportunities of basic education to a child, or of essential health care to the ill, is a failure of social responsibility, but the exact utilization of the educational attainments or of health achievements cannot but be a matter for the personal herself to determine" (Sen, 1999, p. 288).

[11] Scholars of the capability approach often use capability and opportunity without distinction. Here I want to draw a clear line of demarcation between the two concepts used by Sen and by Roemer. With reference to the main differences between equality of capabilities and other equality principles (in particular, equality of resources and equality of well-being) see Granaglia (2007).

advantages or achieved functionings).[12] They both make use of a multidimensional view and assign a major role to individual responsibility, although in different ways. Finally, they both take into account human diversity, focusing on vertical inequalities - whether referring to resources, outcomes or opportunity appropriately defined - as well as on horizontal inequalities (i.e. inequality among groups), as a main source of discrimination. However, the theories differ in their articulation of such inequalities, with no marginal consequences on a methodological level, as well as in terms of policy implications.

The diagram in figure 2 provides a tentative comparison between the two approaches, in a simplified version.

Figure 2 – *Equality of Capabilities versus Equality of Opportunities.*



The observable objects are shown in grey boxes, while the non-observable elements are in dashed ones. Both Sen and Roemer define equality in terms of non-directly observable objects (namely, capabilities and opportunities) and they both

[12] See Chiappero, Grasso, Pareglio (2008) and Ruggeri-Laderchi (2008) for a discussion about the theorising and operating of the capability approach in terms of production functions or as a generating process of well-being. See Chiappero-Martinetti and Salardi (2009) for an econometric estimation of conversion rates.

consider two major sources of inequality. First, individual characteristics which generate unequal personal conditions, whether social, economic or demographic. Second, circumstances experienced by the subjects, which can ultimately create actual discrimination among groups, whether institutional, cultural or political. Drawing on these differences, Roemer defines his "types", while Sen traces back to them the existence of different conversion rates in the transformation process of endowments and resources (both private and public) into functionings.[13]

Here, some clear differences emerge between Roemer and Sen theories. First, Sen considers conversion rates as the result of a process transforming means into ends, but maintains a clear distinction between these three different aspects. Sometimes, Roemer implicitly refers to the conversion process in a way which is not dissimilar from Sen, as in the following passage: (Roemer, 1998, p.6) "Guaranteeing equal per capita financing of educational facilities is, however, non sufficient to provide equal educational achievement, since different children are able to use educational resources (teachers, books, school buildings) with different degrees of effectiveness or efficiency." However, types and circumstances in the end coincide in both Roemer's algorithm and in his empirical studies, where circumstances become a proxy for the (non-observable) opportunities set.

Second, Sen neatly distinguishes resources or endowments from outcomes or functionings. Quite the opposite, in Roemer's empirical studies these two elements (resources and opportunities) tend to overlap. Roemer does not include resources as an independent aspect in determining equality of opportunity. This latter is entirely determined by *circumstance, type, effort, objective, policy* (see Roemer, 1998, 2009) and there is no explicit reference to the amount nor to the nature (public or private) of available resources. Moreover, when defining the "objective" as "the condition for which the acquisition of opportunities is to be equalized", Roemer includes in it things such as wage-earning capacity, although he empirically refers to the available income (Roemer 2009, p. 32). Quite clearly, this double overlap greatly simplifies both formalisation and application of Roemer's approach, allowing him to formulate an algorithm and to operationalize it in a relatively easy manner. However, it suffers from lack of clarity in the arguments and can raise some relevant problems in terms of policy implications as discussed earlier with reference to figure 1.

[13] See Chiappero-Martinetti e Salardi (2009) on the difference between conversion rates and conversion factors.

5. Conclusions

When assessing inequalities, focusing on actual opportunities differs under several respects from assuming outcomes as a focal point, especially when outcomes are conceived only in terms of how income is distributed within a given society. More specifically, approaches focused on opportunities i) acknowledge as morally unfair inequalities due to circumstances independent of individual responsibility, such as gender, race or family backgrounds; ii) attach great importance to individual choice and individual freedom to pursue their own tasks; iii) connect choices and actions to individual responsibility, as long as they are genuinely free (as Sen particularly emphasises).

These three features apply to Roemer's as well as to Sen's approaches. However, the latter sounds more convincing to me, for several reasons which I tried to argue in this paper. First, because it makes a clear distinction between opportunities and circumstances, on one side, and between available resources and achieved outcomes, on the other side. In Roemer's equality of opportunity, these elements tend to overlap, while Sen provides a neater difference between the constitutive elements of his proposal.

These differences are not just a matter of language. As I have tried to point out in this paper, they may have important consequences in terms of policy implications when trying to make up for inequalities.

Aknowledgments

An earlier version of this paper was presented at a workshop in Rome at the CIES (Commissione di Indagine sull'Esclusione Sociale) and at the SIEDS Conference in Florence. I would like to thank the participants and the discussants at these events for helpful discussions and comments. I also thank Elena Granaglia, Renata Targetti Lenti, Italo Magnani, Franco Osculati for their valuable comments and suggestions. The usual disclaimer applies.

References

Alkire S. 2002. *Valuing freedoms: Sen's capability approach and poverty reduction.* Oxford University Press, Oxford.

Arneson R. 1989. *Equality and equal opportunity for welfare.* Philosophical Studies, 56, pp 77-93.

Atkinson T. 2003. *Multidimensional deprivation: contrasting social welfare and counting approaches.* Journal of Economic Inequality, 1, pp 51-65.

Bourguignon F., Chakravarty S.R. 2003. *The measurement of multidimensional poverty*.

Journal of Economic Inequality, 1, pp 25-49.
Bossert W., Fleurbaey M., Van de Gaer D. 1999. *Redistribution, talent and compensation: a second best analysis*. Review of Economic Design, 4, pp 35-55.
Bossert W., D'Ambrosio C., Peragine V. 2007. *Deprivation and social exclusion*. Economica, 74, 296, pp 777-803.
Burchardt T., Vizard P. 2007. *Developing a capability list: Final Recommendations of the Equalities Review Steering Group on Measurement*, CASE Paper n. 121, The Centre for Analysis on Social Exclusion, London School of Economics, London.
Chiappero-Martinetti E., Grasso M., Pareglio S. 2008. *La qualità della vita nelle società avanzate: linee di ricerca per un modello basato sull'approccio delle capacità di Amartya Sen*. Economia e Politica Industriale, 1, 2007, pp 53-73.
Chiappero-Martinetti E., Roche J.M. 2009. *Operationalization of the capability approach, from theory to practice: a review of techniques and empirical applications*. In: Chiappero-Martinetti E. (ed), Debating global society: reach and limits of the capability, approach, pp. 157-200. Fondazione Feltrinelli, Feltrinelli Editore, Milano.
Chiappero-Martinetti E., Salardi P. 2009. *Benessere e fattori di conversione: una stima del processo di benessere a livello individuale*. In: Chiappero-Martinetti E., Pareglio S. (ed), Sviluppo umano sostenibile e qualità della vita. Modelli economici e politiche pubbliche, pp 121-148, Carocci Editore, Roma,.
Cohen G.A. 1989. *On the currency of egalitarian justice*. Ethics, 99, pp 906-944.
Cohen G.A. 2001. *Socialism and equal opportunity*. In: Wolff J., Rosen M. (eds), Political Thought, Oxford University Press, Oxford.
Equality and Human Rights Commission. 2008. *The equality measurement framework*. EHRC Briefing note, London School of Economics, London.
Equalities Review. 2007a. *Fairness and freedom: the final report of the Equalities Review.* www.equalities.gov.uk.
Equalities Review. 2007b. *Consulting for a Capability List*. Research Study conducted by IPSOS-MORI for the Equalities Review, www.equalities.gov.uk
Ferreira F., Gignoux J. 2008. *The measurement of inequality of opportunity: theory and an application to Latin America*. Policy Research Working Paper n. 4659, July.
Fleurbaey M. 1998. *Equality among responsible individuals*. In Laslier J.F., Fleurbaey M., Gravel N., Trannoy A. (eds), Freedom in economics: new perspectives in normative economics, pp 206-234, Routledge, London.
Fleurbaey M. 2005. *Four approaches to equal opportunity*, mimeo.
Fleurbaey M., Maniquet F. 2007. *Help the low-skilled or let the hardworking thrive? A study of fairness in optimal income taxation*. Journal of Public Economic theory, 9, pp 467-500.
Fleurbaey M., Maniquet F. 2009. *Compensation and responsibility*. In Arrow K.J., Sen A.K., Suzumura K. (eds), Handbook of social choice and welfare, pp vol. 2, Amsterdam, North Holland.
Gasper D. 2004. *The Ethics of Development. From Economism to Human Development*. Edinburgh University Press, Edinburgh.
Giovanola B. 2007. *Sulla capacità di essere felici: riflessioni su ricchezza, benessere e libertà a partire dal capability approach*. Meridiana, 56, pp 193-212.

Granaglia E. 2007. *Uguaglianza di capacità: innovazione terminologica o nuova concezione analitica?* QA – Questione Agraria, Rivista dell'Associazione Rossi-Doria, 4, pp 81-105.
Hild M., Voorhoeve A. 2004. *Equality of opportunity and opportunity dominance*, Economics and Philosophy, 20, pp 117-145.
Kolm S.C. 2001. *To each according to her work? Just entitlement from action: desert, merit, responsibility and equal opportunities. A review of John E. Roemer's Equality of Opportunity*. Mimeo, Paris, February.
Kranick L. 1996. *Equitable opportunities: an axiomatic approach*. Journal of Economic Theory, 71, pp 132-147.
Lefranc A., Pisolesi N., Trannoy A. 2006. *Inequality of opportunity vs inequality of outcomes: are western societies all alike?* ECINEQ – Society for the study of economic inequality, Working Papers n.54.
Magni F.S. 2006. *Etica delle capacità.* Il Mulino, Bologna.
Moreno-Ternero J.D. 2007. *On the design of equal-opportunities policies*. Investigaciones econòmicas, vol. XXXI, 3, pp 351-374.
Nussbaum M. 2000. *Women and human development. The capabilities approach*. Cambridge University Press, New York.
Nussbaum M. 2003. *Capabilities as fundamental entitlements: Sen and social justice*. Feminist economics, 9, 2-3, pp 33-59.
Ooghe E., Schokkaert E., Van de Ger D. 2007.*Equality of opportunity versus equality of opportunity sets*. Social Choice and Welfare, 28, pp 209-230.
Ok E.A., Kranick L. 1998. *The measurement of opportunity inequality: a cardinality-based approach.* Social Choice and Welfare 15, pp 263-.287.
Peragine V. 2004. *Measuring and implementing equality of opportunity for income.* Social Choice and Welfare, vol. 22, pp 187-210.
Peragine V. 2005. *Ranking income distributions according to equality of opportunity.*Journal of Economic Inequality, vol. 2, pp 11-30.
Pogge T. 2002. *World poverty and human rights: cosmopolitan responsibilities and reforms*. Polity Press, Cambridge and Blackwell Publisher Ltd, Oxford.
Robeyns I. 2003. *Sen's capability approach and gender inequality: selecting relevant capabilities*. Feminist Economics, 9, 2-3, pp 61-92.
Robeyns I. 2005. *The Capability Approach: A Theoretical Survey*. Journal of Human Development, 2005, 6,1, pp 93-114.
Roemer J. 1998. *Equalities of opportunity*. Harvard University Press, Cambridge MA.
Roemer J. 2002. *Equality of opportunity: a progress report*. Social Choice and Welfare, 19, pp 455-471.
Roemer J. 2009. *Equality: its justification, nature and domain*. Salverda W., Nolan B., Smeeding T., The Oxford Handbook of Economic Inequality, pp 23-39, Oxford University Press.
Roemer J., et al. 2003. *To what extent do fiscal regimes equalize opportunities for income acquisition among citizens?* Journal of Public Economics, 87, pp 539-565.
Ruggeri-Laderchi C. 2008. *Do concepts matter? An empirical investigation of the differences between a capability and a monetary assessment of poverty.* In: Comim F.,

Qizilbash M., Alkire S., The capability approach. Concepts, measures and applications, Cambridge University Press, Cambridge.
Ruiz Castillo J. 2003. *The measurement of inequality of opportunities*. In: Bishop J. and Amiel Y. (eds.) Research in Economic Inequality, 9, pp. 1-34.
Savaglio E., Vannucci S. 2007. Filtral preorders and opportunity inequality. Journal of Economic Theory, vol. 127, pp 474-492.
Schokkaert E. 2007. *Capabilities and satisfaction with life*. Journal of Human Development, 8, 5, pp 415-430.
Saunders P. 1996. *Unequal but fair? A study of class barriers in Britain*. Institute of Economic Affairs, Choice in Welfare Series, n. 28.
Sen A.K. 1985. *Commodities and capabilities*. North Holland, Amsterdam.
Sen A.K. 1987. *The standard of living*. Cambridge University Press, Cambridge.
Sen A.K. 1992. *Inequality reexamined*. Oxford University Press, Oxford.
Sen A. K. 1993. *Capability and well-being*. In: Nussbaum M., Sen A. K. (eds), The Quality of life, pp 31-53, Clarendon Press, Oxford.
Sen A.K. 1995. *Gender inequality and theories of justice*. In: Nussbaum M., Glover J. (eds), Women, culture and development: a study of human capabilities, Clarendon Press, Oxford.
Sen A. K. 1999. *Development as freedom*. Anchor Books, New York.
Sen A.K. 2004. *Elements of a Theory of Human Rights.* Philosophy and Public Affairs, 32, pp 315-356.
Sen A.K. 2009. *The idea of justice*. Allen Lane, Penguin Book, London UK
Van de Gaer D. 1993. *Equality of opportunity and investment in human capital*. Katholic University of Louvain, Leuven.

SUMMARY

In the last decade, the debate on egalitarian justice has increasingly shifted from inequality of outcome to equality of opportunity. This latter paid greater attention to social and economic circumstances which hinder individuals from competing at the same starting level, and to the real opportunities people have to pursue their own life plans and objectives. Amartya Sen and John Roemer are considered the authors who made a significant contribution to the debate, even though they started from different premises, and reached rather different conclusions. In this paper, I will compare and contrast two approaches – equality of opportunity of John Roemer and equality of capability of Amartya Sen - pointing out analogies and differences between these two views.

Enrica CHIAPPERO MARTINETTI, Department of Public Economics, University of Pavia and Institute for Advanced Study, Pavia.

Rivista Italiana di Economia Demografia e Statistica Volume LXIII nn. 1-2 – Gennaio-Giugno 2009

A SURVEY ON DEPRIVATION AND SOCIAL EXCLUSION: ITS THEORY AND LINKS WITH SUBJECTIVE WELL-BEING

Conchita D'Ambrosio

1. Introduction

The survey is based on some of my recent papers joint with Bellani (2009), Bossert (2007), Bossert and Peragine (2007), and Frick (2007a). In a seminal contribution Yitzhaki (1979) proposes to use the Gini coefficient as a measure of relative deprivation and D'ambrosio and Frick (2007a) explore its relationship with subjective well-being as measured by satisfaction with income. A reason for being interested in deprivation is its representation of the degree of discontent or injustice felt by the members of a society. In view of this fact, Podder (1996) criticizes the measures of deprivation proposed in the literature and discusses the reasons why these are unable to capture the phenomenon. Deprivation and inequality are different concepts, hence an index of inequality, such as the Gini coefficient, is inappropriate to measure deprivation. In Podder (1996) the distinction between the two is explained by their relations to envy. " We say that a person i has a feeling of envy towards person j if he prefers to exchange his consumption bundle with that of person j " Podder (1996, p.356). Deprivation is proportional to the feeling of envy towards the better off. Equity---the absence of inequality---is the absence of envy in all economic agents. At the same time, equity coincides with minimum deprivation---all individuals possess the same level of income. In contrast, the upper bounds of deprivation and inequality do not coincide. Maximum inequality is reached when one individual monopolizes the entire total income; maximum deprivation for Podder, on the other hand, is obtained when the society is polarized in two equal-sized groups, those possessing income and those not possessing it.

An analogous distinction with inequality is at the basis of the concept of polarization of Esteban and Ray (1994). The proposed measure of polarization is a variation of the Gini coefficient, where not only alienation pays a role, that is the symmetric gaps of income that are at the heart of the Gini index, but also identification with identical individuals, which is inexistent in the Gini coefficient. Bossert, D'Ambrosio and Peragine (2007) propose an index of deprivation based on the distribution of functioning failures which is a different variation of the Gini coefficient based on the interaction of identification and alienation. The empirical

link between these measures of deprivation and satisfaction with life has been explored by Bellani and D'Ambrosio (2009).

The measures of deprivation look only at incomes or functioning failures in one single time period and do not include information on past experiences of the individuals. Bossert and D'Ambrosio (2007) modify Yitzhaki's index to take into account the part of deprivation generated by an agent's observation that others in its reference group move on to a higher level of income than itself. They formalize an additional idea of Runciman that had not been explored in the literature yet: "The more people a man sees promoted when he is not promoted himself, the more people he may compare himself with in a situation where the comparison will make him feel relatively deprived" (Runciman, 1966, p.19). Relative deprivation of an individual in Bossert and D'Ambrosio (2007) is determined by the interaction of two components: 1) the average gap between the individual's income and the incomes of all individuals richer than it (the traditional way of measuring individual deprivation); 2) a function of the number of people who were ranked below or equal in the previous-period distribution but are above the person under consideration in the current distribution. D'Ambrosio and Frick (2007b) explore the links between subjective well-being, as measured by self declared satisfaction with own income, and the individual's own income history as well as the relative income performance with respect to the others living in the society under analysis.

The remainder of the paper is organized as follows. The measures of deprivation and social exclusion are presented in Section 2. The links with subjective well-being are contained in Section 3. Section 4 concludes.

2. The Deprivation and Social Exclusion Indices

The concept of relative deprivation and its measurement has been introduced in the Economics literature by a seminal paper of Yitzhaki (1979). The definition of relative deprivation adopted is the following: "We can roughly say that [a person] is relatively deprived of X when (i) he does not have X; (ii) he sees some other person or persons, which may include himself at some previous or expected time, as having X, (iii) he wants X, and (iv) he sees it as feasible that he should have X" (Runciman, 1966, p.10). Runciman further adds: "The magnitude of a relative deprivation is the extent of the difference between the desired situation and that of the person desiring it". Following Yitzhaki (1979), income is the object of relative deprivation, as income should be considered an index of the individual's ability to consume commodities. Yitzhaki (1979) suggests to measure individual deprivation as the sum of the gaps between an individual's income and the incomes of all the richer individuals and proposed the absolute Gini as an appropriate index of

aggregate deprivation.

The notation I use in the paper is the following. The sets of all real numbers, all non-negative real numbers and all positive real numbers are denoted by R, R_{+} and R_{++}. Furthermore, N is the set of positive integers. An income distribution is a vector $y=(y_1,...,y_n)\in\mathsf{R}_{+}^{n}$. I indicate the mean of y as $\lambda(y)$. For $y\in\mathsf{R}_{+}^{n}$, $B_i(y)=\{j\in\{1,2,...,n\}\,|\,y_j>y_i\}$ is the set of individuals with a higher income than i.

Yitzhaki (1979), Hey and Lambert (1980) specify the deprivation felt by a person with income y_i with respect to a person with income y_j as:

$$\begin{aligned} d_i(y) &= (y_j - y_i) \quad \text{if } y_i < y_j \\ &= 0 \qquad\quad \text{else} \end{aligned}, \tag{1}$$

while the deprivation function of the person with income y_i is:

$$D_i(y) = \frac{\sum_{j\in B_i(y)} (y_j - y_i)}{n}. \tag{2}$$

Aggregating (2) we obtain total deprivation, which is actually average deprivation, in the whole society:

$$D(y) = \frac{\sum_{i=1}^{n}\sum_{j\in B_i(y)} (y_j - y_i)}{n^2}, \tag{3}$$

which is equal to the product of the mean income $\lambda(y)$ and the Gini coefficient, $G(y)$, i.e. the absolute Gini coefficient.

Following this early literature, Chakravarty (1997), building on Kakwani (1984), proposes to look at a relative concept of deprivation, by taking as a measure of deprivation felt by a person with income y_i with respect to a person with income y_j, their income share differential, $\frac{d_i(y)}{\lambda(y)}$. Now, the total relative deprivation function of the person with income y_i is:

$$D_i^r(y) = \frac{\sum_{j \in B_i(y)} (y_j - y_i)}{n\lambda(y)}. \quad (4)$$

Bossert and D'Ambrosio (2007) modify Yitzhaki's index to take into account the part of deprivation generated by an agent's observation that others in its reference group move on to a higher level of income than itself. A two-period income distribution is a vector

$$(y^0, y^1) = ((y_1^0, \ldots, y_n^0), (y_1^1, \ldots, y_n^1)) \in \mathbb{R}_+^{2n},$$

where y^0 is the income distribution of the previous period and y^1 that of the current period.

An individual measure of deprivation for individual i is a function $BD_i : \mathbb{R}_+^{2n} \to \mathbb{R}_+$. Building on D_i, Bossert and D'Ambrosio (2007) propose the following class of measures BD_i^α, where $\alpha \in [1, \infty)$ is a parameter. For all $(y^0, y^1) \in \mathbb{R}_+^{2n}$,

$$BD_i^\alpha(y^0, y^1) = \alpha^{|B_i(y^1) \setminus B_i(y^0)|} D_i(y^1) = \frac{\alpha^{|B_i(y^1) \setminus B_i(y^0)|}}{n} \sum_{j \in B_i(y^1)} (y_j^1 - y_i^1).$$

Clearly, the Yitzhaki index D_i is obtained for $\alpha = 1$. For higher parameter values, the index assigns weight to the deprivation suffered from the knowledge that others who were previously at or below the income level of i have advanced to a higher income position than i itself. The higher the parameter value chosen, the higher the importance given to being left behind. The dynamic aspect of deprivation depends on the number of those who were at most as rich as i in the previous period but have passed i in the move to the current period. Thus, there is an asymmetry analogous to that present in standard measures of deprivation: only those who passed i matter; their impact on i's deprivation is not counterbalanced by information on those who moved below i. As in the non-dynamic approach, this is the case because deprivation only is being measured and not satisfaction. In the framework of Bossert and D'Ambrosio (2007), individual i would feel satisfied when comparing its income with that of poorer individuals, as in the traditional literature, and would feel even more satisfied with respect to those individuals who used to be richer yesterday and moved to the same level as i or below it in the present period.

Bossert, D'Ambrosio and Peragine (2007) propose a more comprehensive

measure of deprivation considering different aspects of the quality of life of an individual. They assume that, for each individual, there exists a measure of functioning failure, for example the number of them, which indicates the degree to which functionings that are considered relevant in the society under analysis are not available to the agent. The individual functioning failures constitute the primary inputs for the analysis and have been determined at an earlier stage.

For an individual i, $q_i \in \mathsf{R}_+$ is the functioning failure suffered by i in a given period and $\mathsf{B}_i(q) = \{j \in \{1,2,...,n\} \mid q_j < q_i\}$ is the set of individuals with less functioning failures than i. The members of the class of deprivation measures, $BDP_i : \mathsf{R}_+^n \to \mathsf{R}_+$, characterized by Bossert, D'Ambrosio and Peragine (2007) are such that the degree of deprivation for a distribution is obtained as the product of two terms with the following interpretation. The first factor is a multiple of the ratio of the number of agents who are better off than i and the population size. This number is interpreted as an inverse indicator of agent i's capacity to identify with other members of society. The second factor is the average of the differences between q_i and the functioning failures of all agents in $\mathsf{B}_i(q)$. This part captures the aggregate alienation experienced by i with respect to those who are better off. In particular the index is defined by:

$$BDP_i(q) = 0 \quad \text{if} \quad \mathsf{B}_i(q) = \varnothing$$

and

$$BDP_i(q) = \frac{|\mathsf{B}_i(q)|}{n^2} \sum_{j \in \mathsf{B}_i(q)} (q_i - q_j) \quad \text{if } \mathsf{B}_i(q) \neq \varnothing. \tag{5}$$

Although this measure of individual deprivation, reinterpreted in terms of income distributions rather than distributions of functioning failures, resembles that suggested by Yitzhaki (1979), there is an important and substantial difference. Yitzhaki defines only the second factor as the individual deprivation index. Thus, taking into consideration the lack of identification in addition to aggregate alienation is what distinguishes Bossert, D'Ambrosio and Peragine's (2007) approach from earlier contributions. Bossert, D'Ambrosio and Peragine's deprivation measure, reinterpreted in terms of income distributions, resembles that of income polarization suggested by Esteban and Ray (1994). However, it distinguishes itself from the latter in that it is a measure of deprivation where an asymmetry in the alienation component is called for---an individual experiences alienation only with respect to those who are better off. Moreover, a more

comprehensive concept of identification is required because an individual identifies not only with those like it but, instead, with all individuals who are equally well or worse off.

The relativity element of deprivation makes the latter closely related to the concept of social exclusion. Social exclusion depends on the extent to which an individual is able to associate and identify with others. While the concept of deprivation is usually treated as a static concept, social exclusion has important dynamic aspects: an individual can become socially excluded if its condition of deprivation is persistent or worsens over time. Therefore, the measurement of social exclusion requires the inclusion of time as an important variable. Bossert, D'Ambrosio and Peragine's distinction between deprivation and social exclusion is captured by this temporal aspect. An individual experiences a higher degree of social exclusion in situations where deprivation is present in consecutive periods as compared to equal levels of deprivation interrupted by periods without deprivation. Thus, they view social exclusion as chronic relative deprivation in terms of functionings.

Starting from the individual index of deprivation as given by equation (5), Bossert, D'Ambrosio and Peragine account for persistence by giving a higher weight to consecutive periods in a state of deprivation than to isolated ones. A t-period functioning-failure profile involving individual i is a vector $\mathbf{q} = (q^1,\ldots,q^t) \in (\mathsf{R}_+^n)^t$. For all $t \in \mathsf{N}$ and for all $\mathbf{q}$, let $\mathsf{T}_i(\mathbf{q})$ be the set of periods $\tau \in \{1,\ldots,t\}$ such that individual i suffers positive deprivation. For any $t \in \mathsf{N}$ and for any profile $\mathbf{q}$ such that $\mathsf{T}_i(\mathbf{q}) \neq \varnothing$, let $T_i^1(\mathbf{q})$ be the set of consecutive periods beginning with the first period $\tau \in \{1,\ldots,t\}$ such that individual i suffers positive deprivation and ending with the last period $\tau \in \{1,\ldots,t\}$ such that individual i suffers positive deprivation and i is not deprived in the following period if such a period exists; if not, the last period to be included in $T_i^1(\mathbf{q})$ is t. If $\mathsf{T}_i(\mathbf{q}) \cap \{1,\ldots,t\} \setminus T_i^1(\mathbf{q})] \neq \varnothing$, the set $T_i^2(\mathbf{q})$ is obtained from $\{1,\ldots,t\} \setminus T_i^1(\mathbf{q})$ in the same way $T_i^1(\mathbf{q})$ is obtained from $\{1,\ldots,t\}$. Because t is finite, this construction can be repeated until a partition $\{T_i^1(\mathbf{q}),\ldots,T_i^{\ell(\mathbf{q})}(\mathbf{q})\}$ of $\mathsf{T}_i(\mathbf{q})$ is obtained, where $\ell_i(\mathbf{q}) \in \mathsf{N}$ is the number of sets of consecutive periods τ such that individual i suffers positive deprivation.

An individual measure of social exclusion for individual i is a mapping $E_i : (\mathsf{R}_+^n)^t \to \mathsf{R}_+$ that assigns i's level of social exclusion to each profile of intertemporal functioning failures defined as follows:

$$E_i(\mathbf{q}) = 0 \quad if \quad \mathsf{T}_i(\mathbf{q}) = \varnothing$$

and

$$E_i(\mathbf{q}) = \sum_{k=1}^{\ell(\mathbf{q})} |\, T_i^k(\mathbf{q})\,| \sum_{\tau \in T_i^k(\mathbf{q})} \frac{|\mathsf{B}_i(q)|}{n^2} \sum_{j \in \mathsf{B}_i(q^\tau)} (q_i^\tau - q_j^\tau) \quad if \quad \mathsf{T}_i(\mathbf{q}) \neq \varnothing.$$

3. The Relationship with Subjective Well-Being

The relationship between subjective well-being, deprivation and social exclusion using longitudinal data from the German Socio-Economic Panel (henceforth SOEP) have been explored by D'Ambrosio and Frick (2007a, 2007b) and by Bellani and D'Ambrosio (2009). For space constraints I will not include in this survey any review of the findings of D'Ambrosio and Frick (2007b).

Generally, subjective well-being is measured by interviewing people in surveys using a single-occasion, self-report question. Papers on this subject make use of both cross-sectional data (e.g. Eurobarometer Surveys, United States General Social Survey), and panel data (e.g. the German Socio-Economic Panel, the British Household Panel Survey and the European Community Household Panel). D'Ambrosio and Frick (2007a) investigate the relationship between subjective well-being and relative deprivation focussing on panel data since the latter allow to control for otherwise unobserved individual characteristics. This is especially important if these unobservables are systematically correlated with reported subjective well-being. The measure of subjective well-being in the SOEP, i.e. 'satisfaction with income', is measured on an 11-point scale, ranging from 0 ('completely dissatisfied') to 10 ('completely satisfied'). The data used covers the period 1990 (the first data available for the East German sample) to 2004 (the most recent available data when the paper was written). The overall sample contains all adult respondents with valid information on income satisfaction, that is approximately 206,600 observations based on 30,400 individuals in East and West Germany.

The income measure investigated is monthly net household income. This so-called 'income screener' is supposed to give a measure of the more regular income components received by all household members at the time of the interview. This variable might be an inferior measure of economic well-being when compared to annual income since it tends to neglect certain irregular income components (like Christmas bonuses, annual bonuses, etc.) but it certainly fits better to the time-dependent measures of subjective well-being. In order to compare income over time, all income measures are deflated to 2000 prices, also accounting for purchasing power differences between East and West Germany. In order to control

for differences in household size and the economies of scale, an equivalence scale with an elasticity of 0.5, given by the square root of household size, is applied. The measure of relative deprivation used is the modification of the Yitzhaky index proposed by Chakravarty (1997) defined by (4).

Given the ordinal nature of the dependent variable on subjective well-being (perceived satisfaction with income) an appropriate regression model would be an ordered probit. In order to make full use of the panel nature of the data, controlling for otherwise unobserved individual characteristics and potentially different use of the underlying satisfaction scale (running from 0 to 10) across individuals, a fixed effects estimator should be applied. For lack of such fixed-effects ordered probit estimators in standard statistical software packages D'Ambrosio and Frick make use of a fixed-effects regression model, assuming linearity. This assumption appears less problematic given the nature of the dependent variable using an 11-point scale. D'Ambrosio and Frick also run a random-effects model in order to investigate the effects of time invariant control variables, such as gender and migration status.

Correlation results are presented in Table 1 and confirm the findings of Easterlin (2001), suggesting that the natural relationship is more between subjective well-being and relative deprivation rather than between subjective well-being and income itself. The rank is also very highly correlated with subjective well-being, but less so than the relative deprivation measure. The difference between the two measures, rank and relative deprivation, is that the former looks only at the position of the individual in the income scale while the latter takes also into account the distances in incomes. It is worth noting the high, but not perfect, correlation values between equivalent income and relative deprivation (-0.77), and between relative deprivation and rank (-0.96).

Table 1 – *Correlation of subjective well-being, income, income rank, and relative deprivation in Germany 1990-2004.*

	Subjective Well-Being: Income Satisfaction	**Equivalent Income**	**Income Rank**
Equivalent Income	0.3569*	-	-
Income Rank	0.4264*	0.8262*	-
Relative Deprivation	-0.4386*	-0.7729**	-0.9650*

* significant at 1%-level.
Source: D'Ambrosio and Frick (2007a).

Obviously, there is need to investigate whether these relationships hold once we control for various influential factors such as personal and institutional characteristics. In the multivariate regression models, D'Ambrosio and Frick control for sex, age (age squared), marital status, immigration status, education, household composition, homeownership (as a proxy for household wealth) and unemployment. (This unemployment index is calculated at the aggregate household level, relating the number of months in registered unemployment over the previous year to the number of months with potential employment of all adult household members.) It should be noted that in the fixed-effect specification the time independent variables sex and immigration status are dropped from the estimation. In order to control for potential panel or learning effects, a dummy variable identifying individuals with 3 and more interviews as a proxy for the interviewing experience in the panel is included. In order to capture the effect of the state of the economy, D'Ambrosio and Frick include regional unemployment rates at the federal state level. D'Ambrosio and Frick control for the political orientation differentiating individuals with 'strong left', 'left', 'right' , 'strong right', and 'no political orientation' (the latter being the reference group). The political orientation variables are informative of preferences and values of the individuals. 'Lefties' might be more interested in an egalitarian society while 'righties' favour private responsibility and economic success, i.e. 'if you work hard, you also should earn more'. As a consequence of this assumption 'strong righties' *ceteris paribus* should be more satisfyed with higher incomes than others. D'Ambrosio and Frick also present all the models without the political orientation variables as not to influence the main results due to a potential reverse causation between income satisfaction and political orientation. Additional control variables include interaction terms on region (East/West Germany) and year of observation (for readability purposes the latter are not reported in Table 2). First a base model is estimated considering only the above mentioned controls (Models 1 and 6). As a second step are introduced separately in the regression equivalent income as an absolute term (Models 2 and 7), income rank (Models 3 and 8), and relative deprivation (Models 4 and 9). Finally Model 5 (and 10) includes all those measures at once.

Results on the fixed-effects estimators are given in Table 2. Starting with the base model on income deprivation, the personal control variables yield in principle the expected results: more educated and newly married people and those who start to live together with dependent children in the household tend to be more satisfied. There is a significant age effect showing that, controlling for the other covariates, aging decreases satisfaction although at decreasing pace (squared effect). By distinguishing on political orientation D'Ambrosio and Frick conclude that the '(strong) righties' are the only group more satisfied with their income. Homeownership is also positively related to income satisfaction, while the

experience of increasing unemployment within an individual's household has the expected detrimental effect on subjective well-being. The latter can be interpreted as follows: the more the household is affected by unemployment, the less satisfied are all household members with their income. The advantage of the household-based index of unemployment used in the paper is that it also captures an eventual negative effect of unemployment of third household members. In other words, it is not only an individual's own faith of becoming unemployed, which has a detrimental effect on its satisfaction, rather, this is also true if other household members experience unemployment. The institutional control variable also 'behaves' as expected: times of high unemployment exert a dampening effect. With respect to the panel or learning effect, the consistently negative coefficient for 'number of interviews' confirm the findings by Frick, Goebel, Schechtman, Wagner, and Yitzhaki (2006). The interaction terms on region and year of observation (not included in Table 2), *ceteris paribus*, indicate a generally higher level of satisfaction among East Germans.

Table 2 – *Correlates of Subjective well-being (income satisfaction) in Germany 1990-2004 - Results from fixed effects models.*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Income Satisfaction									
Age	-0.057**	-0.069**	-0.060**	-0.056**	-0.057**	-0.057**	-0.069**	-0.060**	-0.056**	-0.057**
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Age squared	0.0001**	0.0001**	0.0001**	0.0001**	0.0001**	0.0001**	0.0001**	0.0001**	0.0001**	0.0001**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Years of education	0.024**	0.014**	0.008+	0.009+	0.008+	0.025**	0.015**	0.009+	0.009+	0.008+
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
# of children in HH	0.042**	0.125**	0.179**	0.159**	0.163**	0.042**	0.125**	0.179**	0.159**	0.163**
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Homeowner	0.118**	-0.001	-0.069**	-0.060**	-0.064**	0.118**	-0.001	-0.069**	-0.060**	-0.064**
	(0.017)	(0.017)	(0.016)	(0.016)	(0.016)	(0.017)	(0.017)	(0.016)	(0.016)	(0.016)
Unemployment Index	-0.010**	-0.008**	-0.007**	-0.006**	-0.006**	-0.010**	-0.008**	-0.007**	-0.006**	-0.006**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Married	0.292**	0.263**	0.237**	0.211**	0.213**	0.292**	0.263**	0.237**	0.211**	0.213**
	(0.019)	(0.018)	(0.018)	(0.018)	(0.018)	(0.019)	(0.018)	(0.018)	(0.018)	(0.018)
3 or more interviews	-0.034*	-0.035*	-0.028+	-0.026	-0.027+	-0.034*	-0.035*	-0.028+	-0.026	-0.027+
	(0.017)	(0.016)	(0.016)	(0.016)	(0.016)	(0.017)	(0.016)	(0.016)	(0.016)	(0.016)
Pol. orientation: Strong Left	0.028	0.035+	0.033+	0.029	0.030	-	-	-	-	-
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	-	-	-	-	-
Pol. orientation: Left	-0.003	-0.006	-0.011	-0.013	-0.012	-	-	-	-	-
	(0.015)	(0.014)	(0.014)	(0.014)	(0.014)	-	-	-	-	-
Pol. orientation: Right	0.028+	0.020	0.016	0.015	0.015	-	-	-	-	-
	(0.017)	(0.017)	(0.017)	(0.016)	(0.016)	-	-	-	-	-
Pol. orientation: Strong Right	0.080**	0.073**	0.067**	0.067**	0.067**	-	-	-	-	-
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	-	-	-	-	-
Regional Unempl. Rate	-0.028**	-0.021**	-0.015**	-0.016**	-0.016**	-0.028**	-0.021**	-0.015**	-0.016**	-0.016**
	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)
Equivalent Income / 1000	-	0.722**	-	-	0.055**	-	0.722**	-	-	0.055**
	-	(0.009)	-	-	(0.014)	-	(0.009)	-	-	(0.014)
Income Rank	-	-	2.442**	-	0.154*	-	-	2.442**	-	0.154*
	-	-	(0.024)	-	(0.075)	-	-	(0.024)	-	(0.075)
Relative Deprivation	-	-	-	-4.104**	-3.716**	-	-	-	-4.104**	-3.716**
	-	-	-	(0.038)	(0.111)	-	-	-	(0.038)	(0.111)
Constant	7.993**	7.388**	6.638**	8.825**	8.614**	8.007**	7.403**	6.651**	8.836**	8.626**
	(0.110)	(0.108)	(0.107)	(0.107)	(0.122)	(0.110)	(0.108)	(0.107)	(0.106)	(0.122)
Observations	206578	206578	206578	206578	206578	206578	206578	206578	206578	206578
Individuals	30396	30396	30396	30396	30396	30396	30396	30396	30396	30396
R-squared	0.03	0.06	0.09	0.09	0.09	0.03	0.06	0.09	0.09	0.09

+ significant at 10%; * significant at 5%; ** significant at 1%. (Standard errors in parentheses).
Additional control variables include interaction terms on region (West/East) and year of observation.

Source: D'Ambrosio and Frick (2007a).

More important to the research question appears to be the comparison of Models 2 (7) to 4 (9), where D'Ambrosio and Frick include alternatively measures of absolute income, income rank, and relative deprivation, respectively. Confirming the bivariate results from Table 1, it appears that after controlling for various personal and institutional characteristics, the highest correlation is given by the relative deprivation. Including income related variables (Models 2 to 5 and 7 to 10) reverses the effect of homeownership on income satisfaction, indicating that net of income effects, new homeowners have higher income aspirations. Including all three measures at the same time (in Model 5 and 10), D'Ambrosio and Frick find only weak relationships for income level and income rank but a large and highly significant coefficient for relative deprivation. These findings suggest that level and changes in subjective well-being in fact are driven more by the relative deprivation an individual derives from its position in a society than by income level itself. However, not surprisingly, a given sum of money, i.e. the absolute level of income, still retains a slightly significant explanatory power for income satisfaction.

The random-effects models (results available upon request) show that women are more satisfied than men, and native born persons are more satisfied than immigrants, in all models.

Bellani and D'Ambrosio (2009) explore the relationship between deprivation and social exclusion proposed by Bossert, D'Ambrosio and Peragine (2007) and satisfaction with life. 'Satisfaction with life' in the SOEP is measured on an 11-point scale, ranging from 0 ('completely dissatisfied') to 10 ('completely satisfied'). The equivalence scale used to transform household monthly income into equivalent income is the modified OECD equivalence scale. Unfortunately the nonmonetary variables which are generally used to calculate deprivation and social exclusion are only available in the SOEP for a limited number of panel years, such as 2000, 2002, 2004 and 2006. These are:

1. living in households that have great difficulties in making ends meet;
2. living in households that are in arrears with (re)payment of housing and/or utility bills;
3. living in households which cannot afford meat, fish or chicken every second day;
4. living in households which cannot afford a week's holiday away from home;
5. living in an accommodation without a bath or shower;
6. living in a dwelling with damp walls, floors, foundations, etc.;
7. living in households which have a shortage of space;

8. not having access to a car due to a lack of financial resources in the household;
9. not having access to a telephone due to a lack of financial resources in the household;
10. meeting their friends or relatives less often than once a month.

Bellani and D'Ambrosio estimate the following reduced-form model of Life Satisfaction (LS_{it}) of individual i in period t:

$$LS_i^t = \theta y_i^t + \phi\lambda\left(y^t\right) + \alpha BDP_i^t + \beta X_i^t + \eta_i + \varepsilon_{it} \quad (6)$$

where y_i^t is the individual i income at time t, $\lambda\left(y^t\right)$ is the mean of the income distribution at time t, BDP_i^t is the individual i deprivation index at time t, and X_i^t is a vector of controls including sex, age (age squared), marital status, education, household composition and unemployment. The individual-specific error, η_i, captures unobserved individual heterogeneity and ε_{it} is an independent error term. This equation is estimated as a linear fixed-effects regression.

Consistent with Easterlin's (2001) finding, simple correlation (Table 3) between subjective well-being and deprivation suggests that this association is much stronger (0.22 on average) than between subjective well-being and income (about 0.12). Bellani and D'Ambrosio first estimate a baseline model, including only individual income and all the contributing factors mentioned above (Model 1). The second specification also includes country specific average income to control for the relative income effect on satisfaction with life (Model 2). As a last step, (in separate regressions), the Yitzhaki (indicated as Yitz in the tables), the Bossert, D'Ambrosio and Peragine (indicated as BDP in the tables) indices of deprivation and the social exclusion index, presented in Section 2 (Models 3, 4 and 5 respectively) are considered.

Table 3 – *Correlation of subjective well-being, income and relative deprivation indices in Germany 2000-2006.*

	Subjective Well-Being: Life Satisfaction	**BDP**	**YITZ**	**Mean Equivalent Income**	**Equivalent Income**
Subjective Well-Being: Life Satisfaction	1				
BDP	-0.2230*	1			
YITZ	-0.2353*	0.9903*	1		
Mean Equivalent Income	-0.0743*	0.0459*	0.0550*	1	
Equivalent Income	0.1222*	-0.3050*	-0.3162*	0.1049*	1

* significant at 1%-level.
Source: Bellani and D'Ambrosio (2009).

In the baseline model (see Table 4) the standard correlates yield results that are in line with many empirical findings in the literature and fairly robust to model specifications. In particular, the results suggest that life satisfaction is significantly associated with income as well as marital status, education and labour market condition.

Not surprisingly, a given sum of money, that is, disposable income, still retains a significant explanatory power for satisfaction, but also the individual relative position matters. The latter negatively affects satisfaction with life confirming the 'relative income hypothesis'. Consistent with earlier work (see, among others, Clark, 2003), becoming unemployed seems to have one of the most important detrimental effects on life satisfaction. Reaching the highest level of education is mildly associated to a lower life satisfaction. Getting married is also significantly associated with an increase in life satisfaction, while getting divorced is associated with a significant decrease. Quite surprisingly having children is associated with a rather small and not significant effect on life satisfaction.

More important for the research question is the comparison between Model 2 and Model 3, when alternative measures of deprivation are considered. If the hypothesis that income and the multidimensional concepts of deprivation are not capturing the same phenomenon holds, then even after the introduction of a nonmonetary index of deprivation as control, both income and the index of deprivation should remain significantly correlated with life satisfaction. This is in fact what is observed in columns 3 and 4 of Table 4 for the Yitzhaki and Bossert,

D'Ambrosio and Peragine indices, respectively. In particular the results indicates that the coefficients for these two indices have the same values. In fact, the introduction of the lack of identification embedded in the Bossert, D'Ambrosio and Peragine index does not seem to have a major impact (the correlation between the indices is approximately 0.99). In Germany alienation, that is the difference between the individual own situation and that of those who are better off, seems to have a dominant effect over lack of identification, that is, the number of people with whom the individual does not identify.

Regarding social exclusion, the results of the linear regression for the last wave of the SOEP (see last column in Table 4) show that persistence in the state of deprivation is significantly negatively associated with life satisfaction. This reinforces the idea of the importance of incorporating also the time dimension in measuring individual well-being.

4 Conclusion

This paper is a survey on measures of deprivation and social exclusion and its links with subjective well-being as measured by self-reported variables of individual satisfaction with income and life. The results show a positive association between the two. Well-being of an individual is not correctly captured by its level of income, as traditional economics modelling assumes. Many other variables play a role and the position in the society the individuals live in also matters.

Acknowledgment

Paper prepared for the XLVI Riunione Scientifica della SIEDS, Firenze 28-30 May 2009. I thank Walter Bossert for comments and all my coauthors for the patience and enthusiasm of working together. Financial support MIUR (Prin 2007) is gratefully acknowledged.

References

Bellani, L. and C. D'Ambrosio, 2009, *Deprivation, Social Exclusion and Subjective Well-Being*, mimeo.

Bossert W. and C. D'Ambrosio, 2007, *Dynamic Measures of Individual Deprivation*, Social Choice and Welfare, vol. 28, pp. 77--88.

Bossert, W., C. D'Ambrosio, and V. Peragine, 2007, *Deprivation and social exclusion*,

Economica, vol. 74, pp. 777--803.
Chakravarty, S.R., 1997, *Relative Deprivation and Satisfaction Orderings*, Keio Economic Studies, vol. 34, pp. 17--31.
Clark, A.E., 2003, *Unemployment as a Social Norm: Psychological Evidence from Panel Data*, Journal of Labor Economics, vol. 21, pp. 323--351.
D'Ambrosio, C. and J.R. Frick, 2007a, *Income Satisfaction and Relative Deprivation: An Empirical Link*, Social Indicators Research, vol. 81, pp. 497--519.
D'Ambrosio, C. and J.R. Frick, 2007b, *Individual Well-Being in a Dynamic Perspective*, IZA Discussion Paper No. 2618.
Easterlin, R.A., 2001, *Income and Happiness: Towards a Unified Theory*, Economic Journal, vol. 111, pp. 465--484.
Esteban, J.-M. and D. Ray, 1994, *On the Measurement of Polarization*, Econometrica, vol. 62, pp. 819--851.
Frick, J.R., J. Goebel, E. Schechtman, G.G. Wagner, and S. Yitzhaki, 2006, *Using Analysis of Gini (ANoGi) for detecting whether two sub-samples represent the same universe: The German Socio-Economic Panel Study (SOEP) experience*, Sociological Methods & Research, vol. 34, pp. 427--468.
Hey, J.D. and P. Lambert, 1980, *Relative Deprivation and the Gini Coefficient: Comment*, Quarterly Journal of Economics, vol. 95, pp. 567--573.
Kakwani, N., 1984, *The Relative Deprivation Curve and Its Applications*, Journal of Business and Economic Statistics, vol.. 2, pp. 384--394.
Podder, N., 1996, *Relative Deprivation, Envy and Economic Inequality*, Kyklos, vol. 3, pp. 353--376.
Runciman, W.G., 1966, *Relative Deprivation and Social Justice*, Routledge, London.

SUMMARY

The survey is based on some of my recent papers joint with Bellani (2009), Bossert (2007), Bossert and Peragine (2007), and Frick (2007a). Starting from the Gini coefficient originally suggested by Yitzhaki (1979) as a measure deprivation, I will introduce other indices of deprivation and social exclusion recently proposed in the literature. Lastly, I will present the relationship between deprivation, social exclusion with subjective well-being as measured by satisfaction with income and life.

Conchita D'AMBROSIO, Associate Professor of Economics, Dipartimento dei Sistemi Giuridici ed Economici, Università degli Studi di Milano-Bicocca, conchita.dambrosio@unimib.it

Rivista Italiana di Economia Demografia e Statistica Volume LXIII nn. 1-2 – Gennaio-Giugno 2009

CROSS REGION AND ECONOMIC PROFILE ITALIAN HOUSEHOLDS WELL-BEING COMPARISONS

Guido Ferrari, Mauro Maltagliati

1. Introduction

"The GDP is dead! Long life for the GDW - From the Gross Domestic Product to the Gross Domestic Welfare". This is how Luigi Biggeri, wittily and happily playing with assonances/dissonances of acronyms, entitled a 2008 seminar held by him in Florence on welfare, quality of life and happiness for Terra Futura.

It is our opinion that this way of seeing things, albeit more generic than the ideas that we intend to present in our communication, could represent the container in which to set our argument and we have made it our own.

And in fact, the issues of household and individual well-being, of poverty, of the quality of life, always the subjects of analysis, discussions and measures, have in recent years attained a crucial relevance in the framework of globalization which has introduced problems that have never before arisen and proposed researches for explanations and solutions, which unfortunately are still far from being found, at least on an acceptable scale.

Some data, resulting from recent researches, gives an alarming measure of the problem even in developed countries.

According to a research by Apiceuropa (2007), even though Europe possesses one of the most advanced social protection systems in the world, there are 68 million European citizens who risk becoming poor and poverty and exclusion are, in accord with Lisbon 2000, challenges that the European governments are trying to confront in a firm and coordinated manner.

The points on which attention is focused are: (i) the perception of poverty, (ii) extreme forms of exclusion and (iii) the perception that citizens possess regarding what is necessary to have a decent standard of living.

Despite the announcement/auspice of the beginning, the GDP continues to be the indicator that is generally used to measure welfare; sometimes income and other macroeconomic indicators are also used, both as substitutes and complements (See, for example, Brandolini, 2008).

Alternatively, there are subjective indicators of evaluation and satisfaction of households relative to the economic situation, the available resources, etc..

It is also possible to construct relations between subjective evaluations of welfare and GDP or relative poverty.

In short, the issue is one of the most studied and debated and the fervour of seminar activity, study meetings, of forums, of awareness initiatives is very much alive, and not only for the social implications that it bears, but also for the political implications that are connected to it.

In effect, at almost all levels of analysis-comment, from the ones more oriented towards the simple description/illustration of the issue to those with more rigorous and refined scientific and methodological contents, debates and contributions are multiplying.

We will limit ourselves to some examples, possibly not even the most significant ones, but which in our opinion can effectively highlight the aforementioned wide spectrum of analyses.

The possible effects of the increase in the price of essential services such as water, electricity and gas on low income households, those which the European Union (EU) prefers to define as " vulnerable consumers", and of the policies that can be put in place for their safeguard to guarantee the access to the consumption of such services have been debated and analyzed in a workshop that took place in Florence (European University Institute, 2008).

To be poor, to end up poorer and to feel poor, and the consequent levels of social exclusion connected to the purely economical fact, are the points analyzed during the two days of study organized by the Ministerial Commission of Investigation in Social Exclusion, "Measures of poverty and policies for social inclusion"(2004) with the aim of formulating proposals and measures to combat this extension of the social-economic vulnerability, in the framework of a combined vision of objective poverty and subjective poverty, of a comparison between the two measures and of a "poverty alarm" that cannot be underestimated.

The ascertainment that an adequate programming of the strategies to contrast poverty and social exclusion implies a correct knowledge of the articulation of such phenomena on the national landscape has led ISTAT to develop the relative statistical information at a regional level, marking such development as one of the objectives of the project "sector-based and territorial statistical information for the structural policies 2001-2008".

With the sample of the Household Budgets Survey (HBS) extended to about 27,000 households and the inclusion of an additional model to collect information regarding situations of deprivation and social degradation that are associated with information on poverty in strictly monetary terms, ISTAT is able to outline a detailed picture of the different aspects that relate to the phenomena of social inclusion/exclusion in Italy.

But we want to immediately reassure that it is not our intention to delve into a close examination of welfare and poverty and analyze the entire debate and the proposals brought forward: we would possibly go far beyond the aims of this convention set up by the "Società Italiana di Economia, Demografia e Statistica" (SIEDS) and, anyway, in a different direction from the one we intend to follow, which is only that of creating a space for reflection and analysis at a sub-national territorial level to illustrate and discuss certain ideas.

In fact, we feel that it could be of certain interest to analyze the welfare of Italian households at the level of large scale territorial divisions, confronting the various situations.

Obviously, to do this we need to confront the income of the households and to make such incomes comparable we have decided to use Equivalence Scales (ES) based on consumption data and on subjective evaluations that the households provide regarding their condition, both sources originating from the investigation into the budget of the Italian households carried out by the Bank of Italy (BI).

In the past we have dealt with, both jointly and separately, ES for comparisons of the households' welfare and cost of children (Ferrari,1999, 2000, 2003; Ferrari-¬Maltagliati, 1999, 2000a, 2000b). But always with an objective approach, based on ISTAT HB data and by using demographically and socially extended complete demand systems.

With the taking place of this SIEDS convention, we decided to re-engage and broaden some considerations on the matter and pursue further results, this time also considering things from a subjective point of view. This was done especially in consideration of the growing interest towards the comparison between objective measures/subjective measures demonstrated by the debate on the matter and for the consequent interesting prospects of reflection and analysis that open up, supported by the availability of evermore reliable subjective data.

In fact, to survey the perception that individuals have regarding the adequacy of their household income to conduct a life which is considered dignified, constitutes one of the central points in the debate over poverty, so much so that we now speak of subjective poverty, a concept of hardship which is broader than those which are traditionally referred to poverty, intended as scarcity of resources, which is to say, in strict economical terms.

According to a ISAE study (2007), the subjective poverty line is notably higher than the one calculated by ISTAT, according to whom the relative poverty line in 2005 was equal, for a household of 2 people, to 936 Euros a month(with a portion of poor households equal to 11%), whilst according to ISAE, such threshold was equivalent to 1800 Euros a month (with a portion of poor households equal to 74%).

Even though in the aforementioned study there is a reference to a territorial repartition, underlining that the incidence of subjective poverty is higher in the southern regions and with singles and couples, and with lower income households, there is no detailed analysis on a sub-area level.

We also believe that there are no other studies, carried out with objective or subjective methods, regarding the comparison of welfare in Italian households under territorial divisions and economic profiles, at least as far as we know.

There are some studies at a regional level: apart from the aforementioned investigation by ISTAT on poverty and social exclusion in Italian regions, in which estimates on regional poverty are provided, we will quote the work of Ferrari (2004) in which ES are calculated using ISTAT HB data based on demand systems extended in both a demographical and social sense , with a reference family of two members, between households classified in expense groups (income), working condition of the head of the family and composition of the household itself, both for all the Italian regions compared to Italy and for some regions compared to others.

Our work will therefore be articulated on an initial, brief illustration of the methodology that supports the objective and subjective ES that we will employ to equalize the household incomes. We will then pass to the estimation of the ES and to the subsequent one of equivalent incomes, subjective and objective, the latter at the level of the five classical territorial divisions (North East, North West, Centre, South, Islands) and for Italy as a whole. These equivalent incomes will be used to compare household welfare under territorial divisions. A comparative analysis of the results will close the work.

2. The equivalence scales employed

The objective ES on which we have based our work are those that result from the Working-Leser model, from the name of the economists that first proposed them[1], and which are probably the ones which have been most used[2] in empirical applications (even though obviously they are not the only conceivable ones[3]), which are based on the Engel approach:

$$w_n = a + b \cdot log(S) + c \cdot (N_c - N_r), \qquad (1)$$

where:

w_n is the expenditure share for necessary goods, those for which the fraction of expenditure decreases with the increase of income and of which food is only a part, S the total income-expense, N_c the number of the generic household members, N_r the number of the reference household members (2 members), *a, b* and *c* are parameters.

There are two reasons for the "success" of such form (and of its extensions): good adaptability to empirical data and relative simplicity in estimating the parameters 4.

In the model the number of household members is the (only) demographic variable employed. Other choices are possible: to consider, besides the number of household members, also other characteristics such as, for example, sex and age, depends on the aim of the model.

Regarding the signs of the parameters, *b* will be negative and *c* positive based on what previously said. As for parameter *a,* it too will be positive, because it represents the fraction of expenditure of the reference household with *log(S)=0* or *S=15*.

In almost all the applications that we are aware of, Engel's principle has been applied only to the food category. In recent years some attempts have been made to extend Engel's principle to other categories of goods, up to the inclusion of all necessary goods 6. In reality though, in our opinion, the so called "classical" approach is still to be preferred for an essential reason. As can be verified by using very detailed *databases* 7, the expenses for necessary goods are composed almost entirely of foodstuffs and household expenses, which include rent (or the rental value if the house is owned) and energy (heating and electricity). And the household expenses come under the *collective consumption* services, the amount of which does not vary (or varies slightly) with the variation in the number of household members. For this reason the percentage of household expenses cannot be used to compare the welfare of households with differing incomes and/or household members.

In this work, for the calculation of the ES we therefore only consider the fraction w_a of food expenses:

$$w_a = a + b \cdot \log(S) + c \cdot (N_c - N_r) \qquad (2)$$

The parameters of this equation, *a, b* and *c,* can be estimated with the Least Squares method, if we possess the data relative to the variables w_a , S and N for a sufficiently large sample of households.

Let us suppose, once we have estimated such parameters, that we want to calculate the ES of a household with N_c members, after having conventionally established that the reference household is composed of $N_r = 2$ members.

We will have for the two households:

$$w_{ac} = a + b \cdot log(S_c) + c(N_c - 2) \qquad (3)$$

$$w_{ar} = a + b \cdot log(S_r), \qquad (4)$$

where S_c and S_r are the expenses sustained by the household *c* and *r* (reference), to reach, respectively, the food expenditure amounts w_{ac} and w_{ar}.

Subtracting the second equation from the first we will have, for the same level of food expense amount(or the same level of material welfare):

$$b \cdot log\left(\frac{S_c}{S_r}\right) + c \cdot (N_c - 2) = 0, \qquad (5)$$

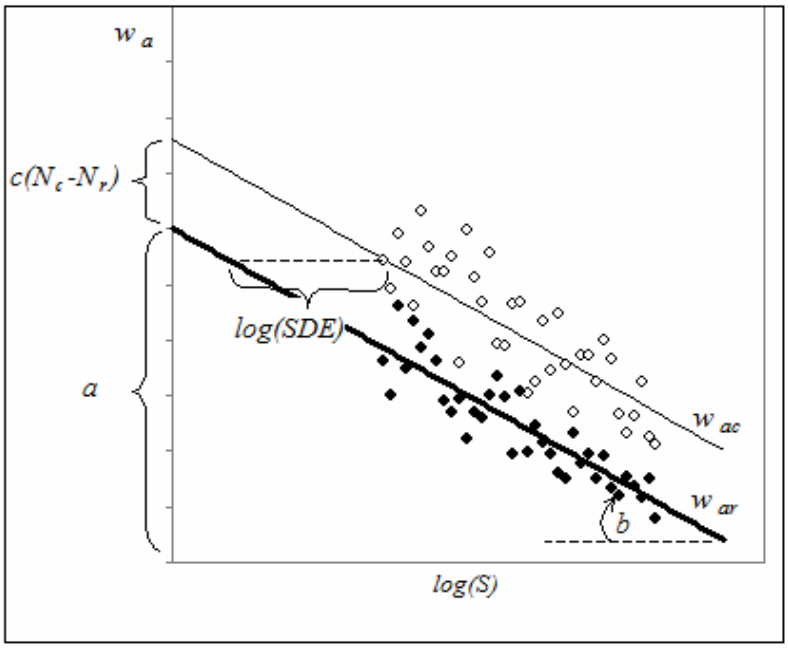
and so,

$$SDE = \frac{S_c}{S_r} = e^{[-c \cdot (N_C - N_r)/b]} \qquad (6)$$

This relationship represents the sought ES.

As can be seen in Fig. 1, the curves that represent the fraction of food expenses are both linear and parallel (both with slope equal to *b*). In practice, the curve w_{ac} is simply the curve w_{an} "shifted upwards" by a factor of $c(N_c-2)$. This has an important consequence: because the w_a are two straight lines, their "horizontal distance" is also constant. So the value $log(S_c/S_r)$ is constant (for any level of *log(S)*, as indicated by the equation (6)), and so the ES is constant. In economical terms, this means that for any level of income-expenditure, the percentage increase in the cost which is necessary for a generic household to have a standard of living equal to the reference household, does not depend on the economical welfare level of the latter.

Fig. 1 – *Food expenditure fractions according to the Working-Leser model.*



This property is called ESE (*Equivalence Scale Exactness*) property.

As for the subjective ES, they are constructed by us according to the following simple reasoning: after fixing a level of welfare, we consider all the households that declare having such level and we divide them in groups according to the number of their members. We then calculate the average income for each group and we divide it by the average income of the reference group (2 members) that has declared the same level of welfare.

We are then able to obtain the ES for the welfare level considered and for that number of members.

3. The Equivalence Scales and Equivalent Incomes Estimation

Using the responses to the questions relating to the food purchases of the Italian households budget sample survey of 2006 (about 7,500 households) carried out by the BI, which were as follows:

What were the average monthly expenses for food purchases? Considering the expenses for foods in supermarkets or similar shops and the expenses for meals regularly consumed in restaurants, cafes and eateries

We obtained the following objective ES

Tab. 1 – *Objective ES (Working-Leser model).*

Household members	1	2	3	4	5
ES	0,768	1	1,301	1,693	2,203

To construct subjective ES we then considered the answers to another two questions, from the same BI survey, regarding the subjective perception of the well-being by the head of the household, which were:

In your opinion, how *much does a household like yours need to live without luxury but without giving up the essentials?*

And

The income available allows your household to arrive to the end of the month:
-with many difficulties 1
-with difficulties 2
-with a few difficulties 3
-quite easily 4
-easily 5
-very easily 6

(Because the answer number 6 was given by a very small number of individuals, we decided to include them with the ones of answer 5).

The answers to the second question allowed us to divide the households into 5 groups of "subjective well-being": poor (answer 1); almost poor (answer 2); normal (answer 3); almost rich (answer 4); rich (answers 5 and 6).

We can immediately observe (as can be seen in Graph 1 that follows), even before the evaluation of subjective ES, that the types of households that "perceive" a better welfare are those of 2, 3 and 4 members in the North and the Centre.

More specifically, it is the North East that displays the better situation, with 14, 15 and 16 % of " poor and almost poor" respectively in households with 2, 3 and 4 members. In addition, the North West scores percentages which are just slightly higher for households with 2 and 4 members, and equal percentages for households with 3 members.

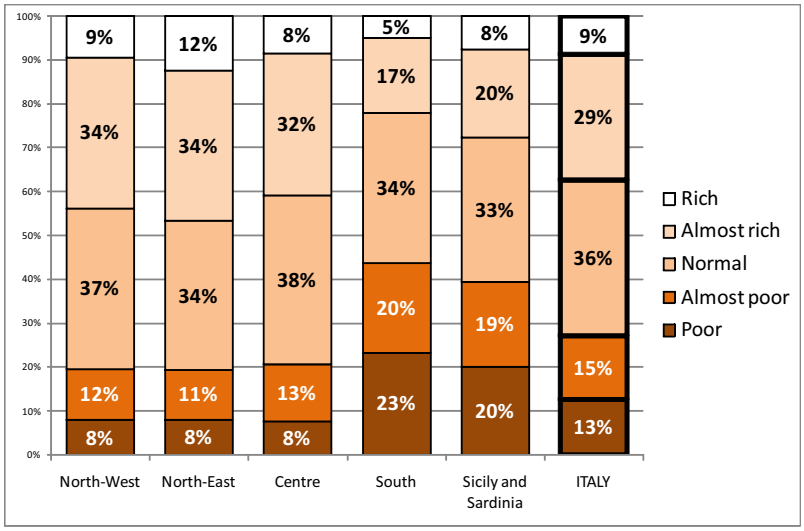
As a whole, the fraction of households that feel almost poor or poor never exceeds 18%.

On the other hand, the perception of being rich or almost rich is observed in 54% of households with 2 members and in 50% and 51% of households with 3 and 4 members in the North East, and by 47, 50 and 43% of households with 2, 3 and 4 members in the North West.

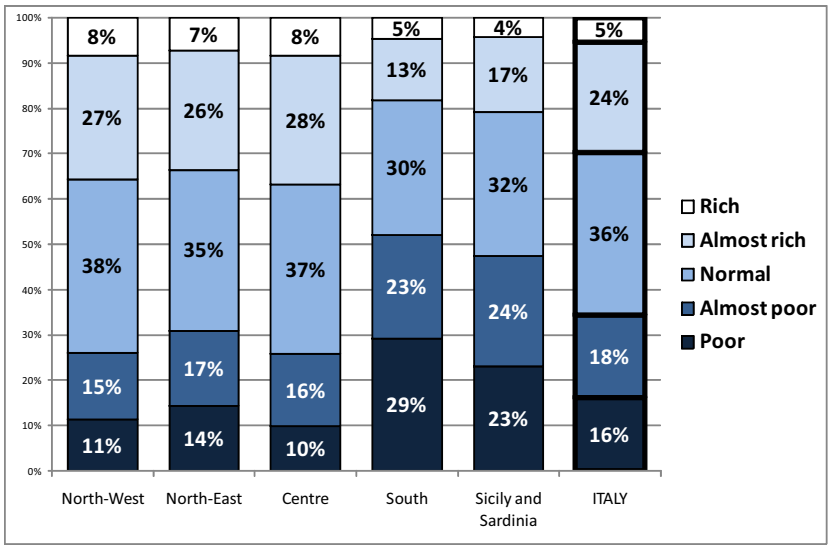
In the case of the Centre, the households that declare feeling almost poor or poor add up to 21,19 and 15% in households with 2, 3 and 4 members, which, as we can observe, are percentages only slightly higher than those in the North. At the same time, households that "are better off" (almost rich or rich) represent respectively 41, 45 and 40% in households with 2, 3 and 4 members, percentages these with are more markedly inferior to those of the North. And in fact, it is the amounts of "normal" households that in the Centre are a lot higher than those in the North (an average of about 40% against an average of about 35%).

Graph. 1 – *Subjective perception of well-being by number of household members and territorial division.*

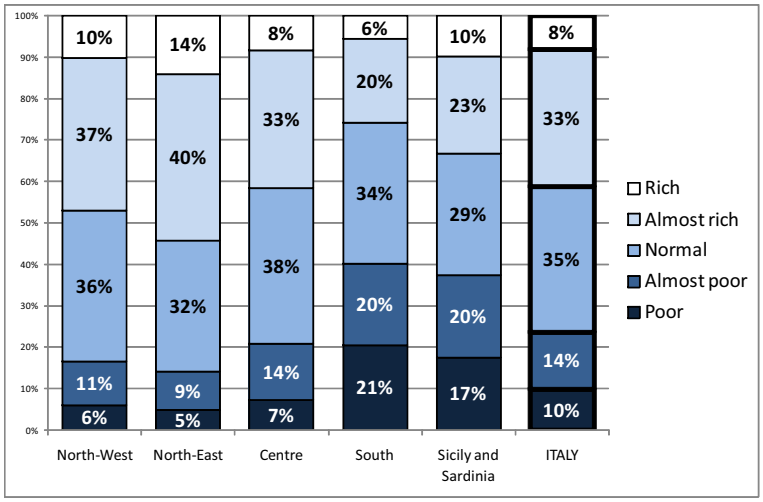
All households



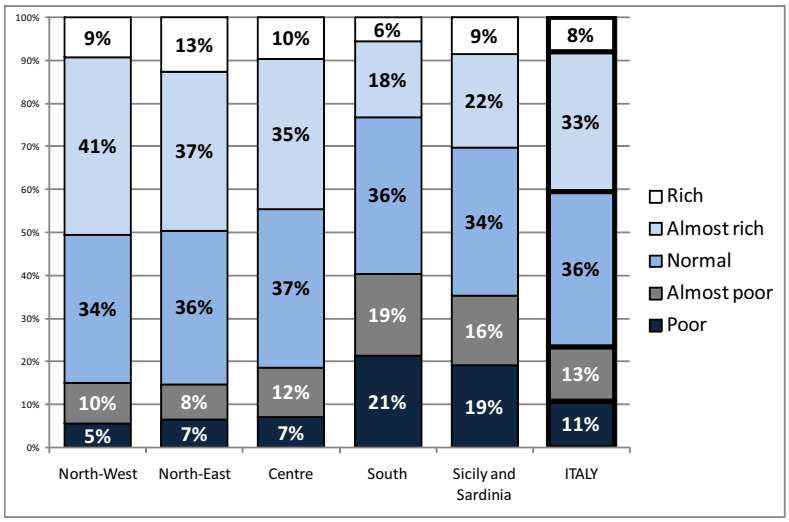
Graph. 1 (cont.) – 1 Member



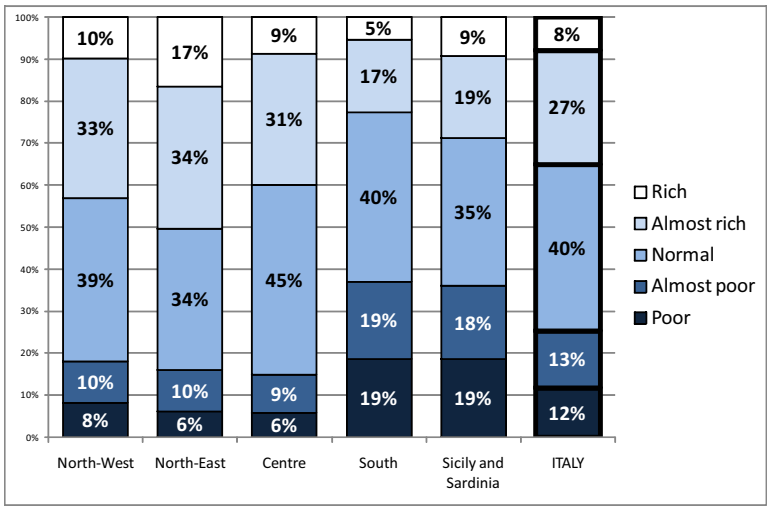
Graph. 1 (cont.) – 2 Members



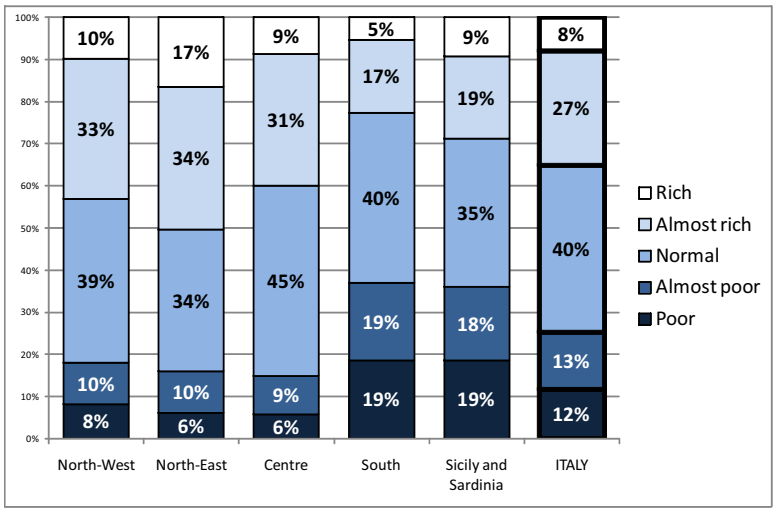
Graph. 1 (cont.) – 3 Members



Graph. 1 (cont.) – 4 Members



Graph. 1 (cont.) – 5 Members



The households which are "worst off" are those composed by only one and five household members in the South, where respectively 52% and 55% of households feels poor or almost poor, with households which are "better off" accounting for 43% in both cases. This is a consistent fraction, equal to the one for the Centre, because of the amount of normal households which accounts for only 23% both in households with one member and households with 5 members.

The islands follow the South at a very close distance, with a percentage of poor or almost poor equal to 47% of households with one member and 44% of households with 5 members.

The Italian average reproduces the same conditions: the minority households which are "worst off"(24, 23 and 25% respectively) are those of 2,3 and 4 members, whereas those that are "better off" are still the same, with 33, 33 and 27% respectively. The households with only one member that declare themselves as "feeling worst off" represent 34% (compared to 24% that declare themselves as "feeling better off") , as opposed to households with 5 members that consider themselves as being "worst off" in 39% of the cases and "better off" in 23% of the cases.

If we don't consider the number of household members, the households which are "worst off" represent 19% in the North East and 20% in the North West, 21% in the Centre, 43% in the South and 39% on the Islands; those which are "better off" amount to 46% in the North East, 43% in the North West, 40% in the Centre,

22% in the South and 28% on the Islands. All the fractions, these, which reproduce, albeit with a few approximations, the situation on a household composition level.

For Italy as a whole, we can observe that the households which are "worst off" add up to 28% and those which are "better off" add up to 38%, whilst the normal ones add up to 36%.

Let us now return to the subjective ES.

The following table shows the average monthly household incomes according to number of household members and subjective perception of well-being.

Tab. 2 – *Average monthly household incomes (in Euros) according to number of household members and level of Subjective Well-being Perception (SWP). (Year 2006.Euro).*

	Number of household members				
SWP LEVEL	1	2	3	4	5
Poor (1)	943	1253	1722	1660	1898
Almost poor (2)	1191	1699	2172	2350	2375
Normal (3)	1489	2163	2775	2931	3044
Almost rich (4)	2150	3073	3859	4390	4542
Rich (5)	3526	4552	5356	5634	6472

By dividing all the values in the columns by those in column 2 we obtain the subjective ES that can be found in Tab. 3.

We then considered all the monetary incomes and, for each household composition, we divided them by objective ES, obtaining the equivalent incomes according to the objective approach.

The same monetary incomes, for each household composition and for each subjective welfare group, were then divided by the subjective ES, obtaining the equivalent incomes according to the subjective approach.

Tab. 3 – *ES according to number of household members and SWP levels.*

	Number of household members				
SWP LEVEL	1	2	3	4	5
Poor (1)	0,753	1,000	1,374	1,325	1,515
Almost poor (2)	0,701	1,000	1,278	1,383	1,398
Normal (3)	0,688	1,000	1,283	1,355	1,407
Almost rich (4)	0,700	1,000	1,256	1,429	1,478
Rich (5)	0,775	1,000	1,177	1,238	1,422

In both cases, we divided the equivalent incomes into 10 groups of equal number of household members and then calculated the arithmetical mean of the incomes, as shown in Tab. 4 and 5.

Tab. 4 – *Equivalent average monthly household incomes (in Euro) according to an objective approach and monetary incomes.*

		Average monthly monetary income by number of household members				
group	Average equivalent income based on objective ES	1	2	3	4	5
1	664	532	634	827	1145	1441
2	1081	828	1086	1401	1838	2356
3	1357	1040	1359	1773	2291	2996
4	1593	1217	1608	2073	2682	3550
5	1834	1407	1838	2393	3090	4055
6	2088	1616	2085	2715	3510	4591
7	2380	1833	2373	3100	4043	5178
8	2751	2108	2753	3595	4632	6094
9	3322	2532	3340	4320	5596	7354
10	5628	5049	5475	6712	10120	10564

The equivalent incomes of these tables clearly depend on the ES with which the monetary incomes have been converted.

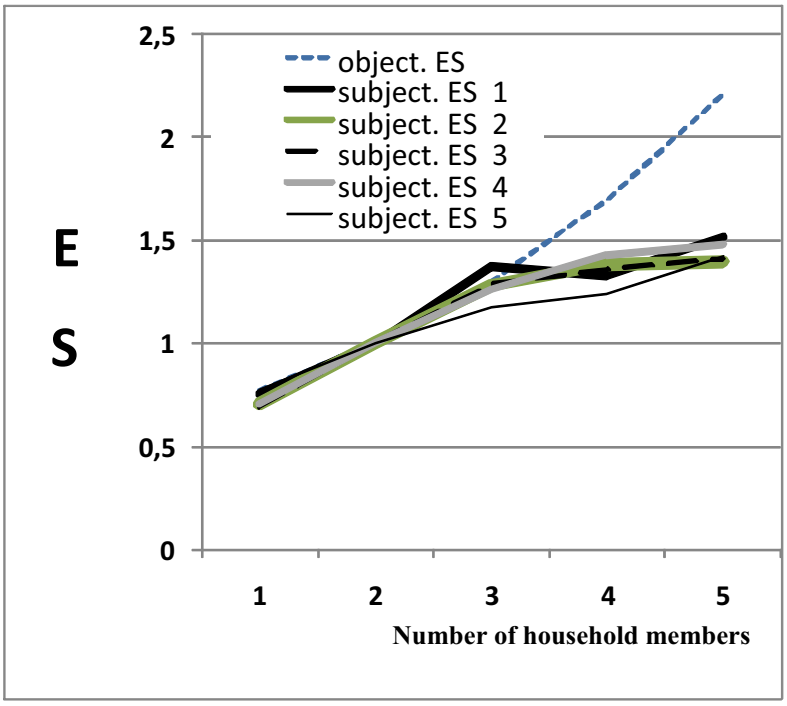
Such scales, in the objective approach, are first of all global and do not consider the differences in the level of income (to simplify, between rich and poor) and secondly, because the logarithm grows linearly as the household member numbers increases, they increase more than proportionately. In other words, not only there are no "returns to scale", but there actually are "diseconomies of scale".

In Fig. 2 we can observe the more than proportional increasing trend of the objective ES and the less than proportional increasing trend of the 5 subjective ES, as the number of household members increases.

Tab. 5 – *Equivalent average household monetary incomes (in Euro) according to the subjective approach and the monetary incomes.*

		Average monthly monetary income by number of household members				
group	Average equivalent income based on subjective ES	1	2	3	4	5
1	729	641	725	819	875	990
2	1173	1007	1183	1362	1395	1465
3	1451	1234	1450	1696	1739	1780
4	1716	1455	1724	1969	1992	2044
5	1999	1683	2002	2290	2310	2408
6	2305	1920	2303	2625	2674	2745
7	2665	2202	2669	3009	3085	3103
8	3124	2562	3131	3489	3562	3594
9	3809	3141	3798	4238	4316	4403
10	6545	6853	6367	6539	7384	6780

Fig. 2 – *Graphical representation of subjective and objective ES.*



We have ordered the equivalent incomes of all the households, independently on the number of household members, by income increasing and then divided them in ten equally numerous groups, indicating them with a progressive numbering from 1 (first group of households, the poorest, delimited above by the first decile) to 10 (tenth group of households, the richest, delimited above by the tenth decile). The deciles based on the two approaches are as follows:

Tab. 6 – *Equivalent income deciles in the objective and subjective approaches.*

Deciles (Euro)	I	II	III	IV	V	VI	VII	VIII	IX
Objective Approach	920	1229	1479	1713	1957	2225	2552	2989	3730
Subjective Approach	1009	1328	1586	1857	2145	2476	2871	3417	4311

We then took the 5 geographical areas, considering for each of them the percentage of households belonging to each of the 10 groups and comparing it with the national percentage. The positive or negative differences in the percentages of each group give us a measure of the higher or lower level of welfare in the geographical area compared to the national average.

In this way we have obtained the graphic representations that can be found in Graph 2 and 3, respectively for the objective and subjective approach.

We have preliminarily determined a poverty line considering as a reference, as done by ISTAT, the average income of a household with one member. In our case it amounts to 1,375 Euro a month. So all households with an equivalent income below this threshold are poor. In the increasing distribution of equivalent incomes obtained both with objective ES and subjective ES, this figure is almost equivalent to the second decile. As a consequence, the households which lie below the second income decile are poor; in other terms still, 20% of households is poor.

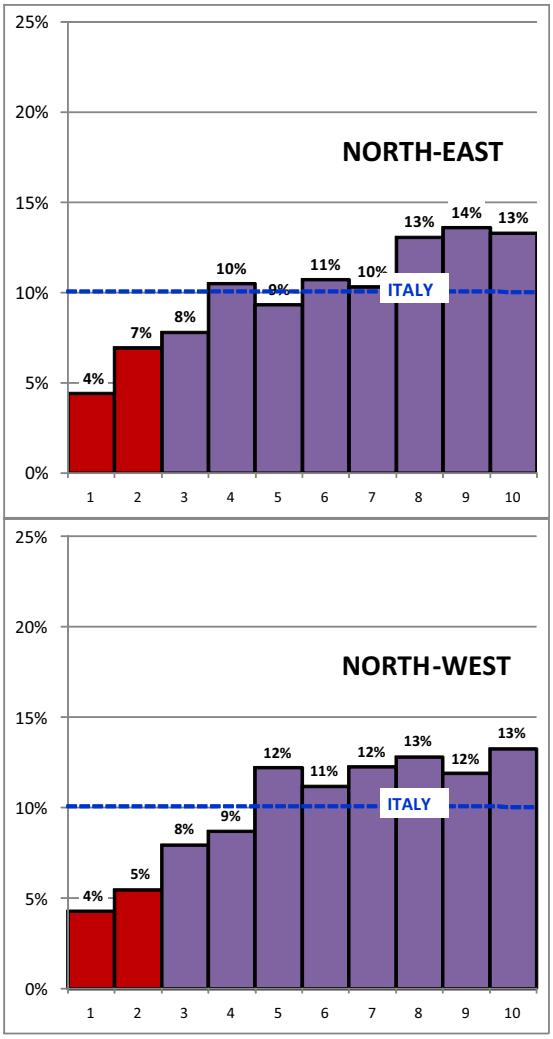
If we consider the objective approach we can see that the North West, the North East and the Centre (the first three frames of Graph 2) have a fairly uniform behaviour in broad terms: poor households represent 9% in the first case and 11% in the other two. So, in these areas the poor are clearly less than the national average.

If we imagine for convenience, with approximation and simplifying, that the "extremely poor" households fall below the decile, we can see that the 3 geographical areas are particularly favoured, with percentages of 4 and 5 %, clearly below the 10% national average.

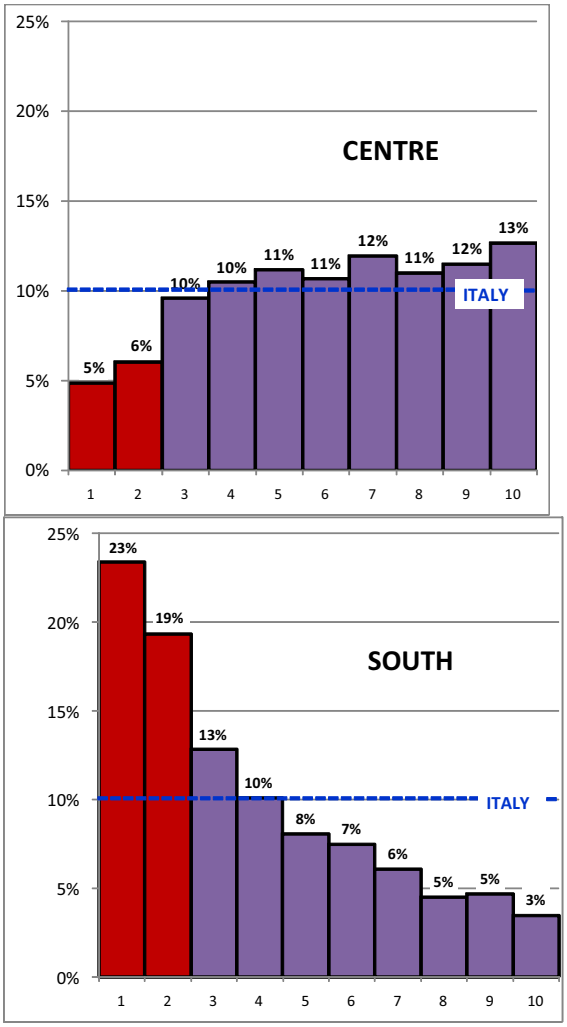
On the other hand, the "extremely rich" (that we can imagine above the IX decile) are 3 percentage points above the national average.

Things are very different in the South and the Islands. In the South, poor households add up to 42%, more than double the national average, whilst on the Islands the percentage clearly decreases, reaching 36%. The extremely rich are 3% in both geographical areas, a percentage which is clearly below the national average.

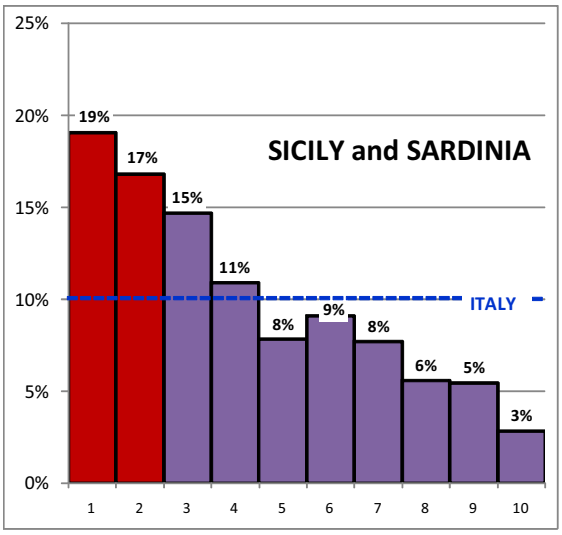
Graph. 2 – *Comparison between equivalent incomes by deciles and geographical areas. Objective approach.*



Graph. 2 (cont.)


25%
20%
15%
10%
5%
0%
CENTRE
5%
6%
10%
10%
11%
11%
12%
11%
12%
13%
ITALY
1
2
3
4
5
6
7
8
9
10
SOUTH
23%
19%
13%
10%
8%
7%
6%
5%
5%
3%
ITALY

Graph. 2 (cont.)



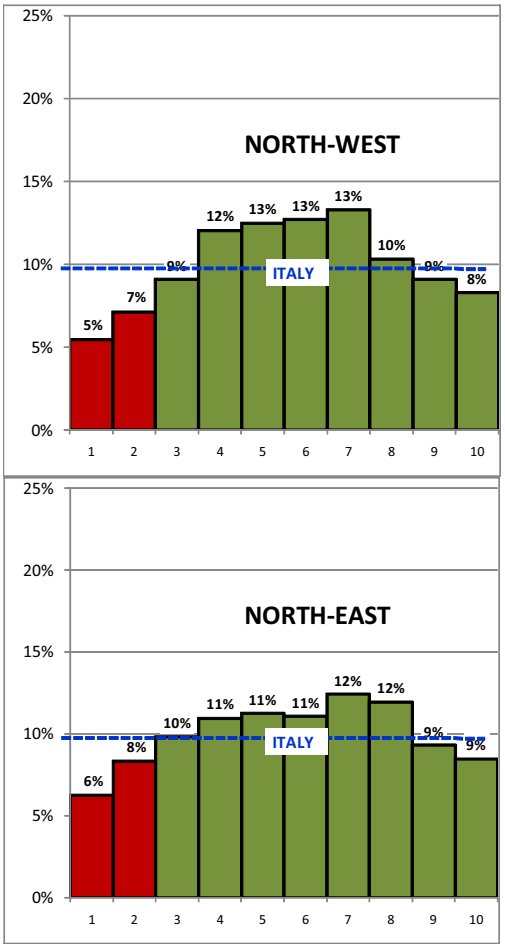
A quick look at the situation resulting from the subjective approach (Graph 3) allows us to see that poor households are 12% in the North West and 14% in the North East and the Centre. We are still well below the national average of 20%, but the figures are above those highlighted by the objective approach.

The "extremely poor" add up to 5% in the North West and to 6% in the other two geographical divisions. Again in this case, the percentages are slightly above the ones highlighted by the objective approach.

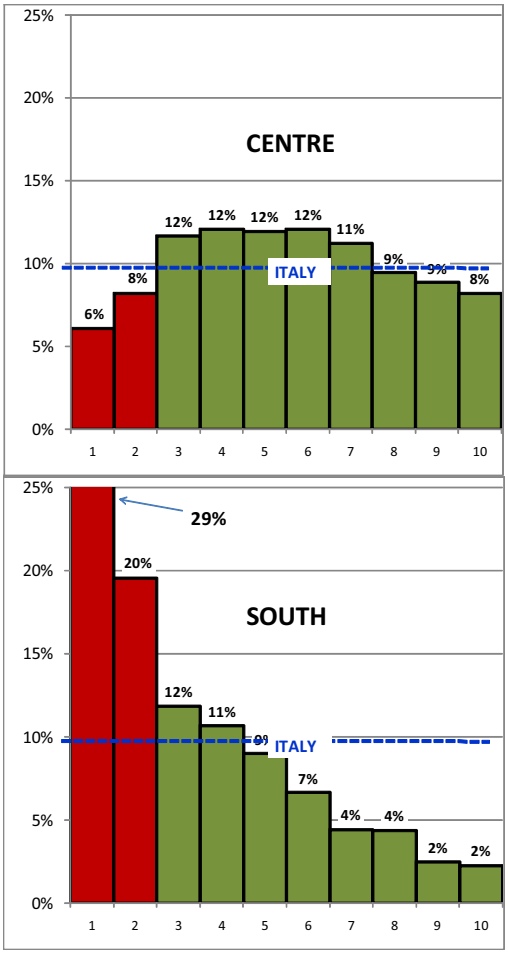
Once again, the South and the Islands display, in this approach too, a totally different situation: in the South 45% of households are poor; in the Islands this percentage decreases slightly, reaching 41%. They represent very high figures, even higher than the ones highlighted by the objective approach.

The "extremely rich" are very few: 2% of households in both geographical areas.

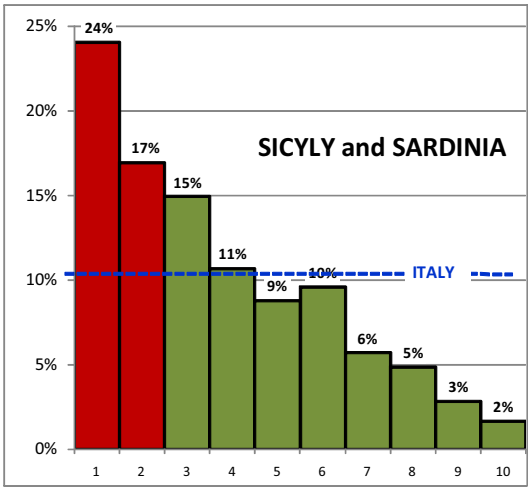
Graph. 3 – *Comparison between equivalent incomes by deciles and geographical areas. Subjective approach.*



Graph. 3 (cont.)


25%
20%
15%
10%
5%
0%
CENTRE
6%
8%
12%
12%
12%
12%
11%
9%
9%
8%
ITALY
1
2
3
4
5
6
7
8
9
10
25%
20%
15%
10%
5%
0%
SOUTH
29%
20%
12%
11%
9%
7%
4%
4%
2%
2%
ITALY
1
2
3
4
5
6
7
8
9
10

Graph. 3 (cont.)



4. Conclusions

In the context of the reflections referring to the comparative analysis of the households' well-being welfare for the 5 large Italian geographical areas that starts with the equivalent incomes estimated both with a subjective and an objective approach, that we announced at the beginning, some interesting evidence has emerged, some already well defined and conclusive, some other still fuzzy and claiming for further deepening and cognitive contributions.

The cue to further expand our reasoning and carry out empirical verifications regarding poverty and welfare was given to us by the persistence of the absolute relevance of the measure of welfare and poverty and, actually, by its increase in the context of globalised societies and economies, and by the fact that the attention has shifted in an ever increasing way towards the perception that households have of poverty and of how much it is considered by them necessary to have an acceptable standard of living, as well as towards social and economical exclusion. In essence, on a subjective view of welfare and poverty.

From this, our interest for a broadening of the researches already carried out in an objective way and of a comparison with the results deriving from a subjective approach. All the more as the available "perception" data is by far more reliable than in the past and we have at our disposal the results of the sample investigation carried out by the BI. If we then can or must extend the analysis and comparison also to evidence deriving from other subjective sources(for example the data from

the multi-scope survey by ISTAT or from public research institutes others than the BI) and/or from other objective contexts (for example, through the use of more complex models, such as the complete demand systems, but on subjective data) is a totally reasonable and understandable idea that we intend to examine more closely in future research.

We have estimated subjective and objective ES, the formers for 5 levels of SWP. It appears to us that in both cases the results that we have obtained are in line with what has been highlighted in similar researches, and so, very strongly founded.

We have empirically confirmed the differential characteristic of the objective ES compared to the subjective ones; that the formers are much more "rigid" in construction than the latter ones, which instead allow "returns to scale" in relation to the household dimension. The subjective ES also have the advantage, obvious for what has just been said and for their intrinsic nature, of being possible for levels of SWP.

Having extrapolated for our aims three questions from the group of those asked in the consumption survey by the BI, we were able to carry out a sort of preliminary comparative analysis on the perceptions of households based on their monetary incomes, introducing also the spatial context, by taking the analysis to a geographical area level, which gave to us some confirmations and a few small integrations.

In this way, households which perceive a higher level of well-being, that feel " better off", are the "average" ones, with 2, 3 and 4 household members in the North and Centre of Italy. Evidence this that confirms the expectations deriving from a conviction dictated by almost all information sources, such as television, radio, newspapers, etc. and reinforced by the observation that the North East is the area that displays the best situation.

Households that feel "worst off" are those in the South and with 5 household members- and this is a result that could be expected- and 1 household member, a result, this, which appears slightly more difficult to explain.

As for the Islands, the situation is very similar to that of the South, with only a few slight differences.

Equivalent incomes, for a household with two members, obtained in both cases, are such, as has been confirmed by the analysis that followed, that we have to assume there are no relevant contradictions between the type and economic quality of the life that it lives and that is represented by the structure and qualitative level of its expenses and the perception of its own well-being situation that a household has. This seems to us a statement worth paying attention to, that can clear the way for further in depth investigations.

The comparative analysis of equivalent incomes has been carried out by preliminarily fixing a poverty threshold established by the average income of a 1

member household, equal to 1,375 Euros a month, a value that, in our equivalent income distribution, both objective and subjective, falls close to the second decile.

On this base, and considering the objective equivalent incomes, we have observed that in the North and Centre of Italy, poor households represent about 10% of households, significantly below (exactly half) the national average. If, with a bit of forcing and approximation, we consider as very poor the households of the first national decile, we can see that extremely poor households, in those same geographical areas, are significantly below the national average of 10%, only reaching 4%.

In this case too, as was to be expected, the South and the Islands display a completely opposite tendency compared to the former.

In the South, poor households are more than 40%, twice the national average; in the Islands things are slightly less negative, with 36% of poor households.

The comparison of these results with those that come from subjective equivalent incomes, shows that the situation in the North and the Centre of the country, despite being far better than the national average, is slightly less "good", with poor households accounting for more than half of the national average.

The South and the Islands are still by far the most penalised areas and it is here that we can observe percentages of poor households constantly higher than those obtained with the objective approach.

We must underline the issue that, despite the fact that a few differences between the objective and subjective approach may appear as being not quite negligible, we can state with reasonable conviction that the two approaches do not lead to significantly different results. Naturally, whilst waiting for further comparative analyses, especially useful because our research constitutes only the first step in a process that requires further tests and confirmations.

As we have seen, the comparative analysis on a geographical division level was carried out only for the households as a whole and not for households divided by number of members. This aspect, that has nonetheless granted us important observations and comparative conclusions between the geographical divisions (and has also granted us comparisons between the evidence of the subjective and objective approaches) is without a doubt a limit, seeing as a more comprehensive area analysis that would include the household structure would surely be more informative. Lack of time and space has led us to limit ourselves for the moment to a more "aggregated" analysis, but it is our intention to proceed immediately to the expansion and development of the work in the sense of considering also the number of household members.

References

Apiceuropa (2007), Una ricerca sulle percezioni della povertà, http://www.apiceuropa.com.
Brandolini, A. (2008), On Applying Synthetic Indices of Multidimensional Well-being: Health and Income Inequalities in Selected EU Countries, *Temi di Discussione (Working Papers)*, Banca d'Italia, N. 668.
Carbonaro, G. (1985), Nota sulla scala di equivalenza, Presidenza del Consiglio dei Ministri, Commissione di indagine sulla povertà, "La povertà in Italia – Studi di base", Roma.
Commissione Ministeriale di Indagine sull'Esclusione Sociale, "Misure della povertà e politiche per l'inclusione sociale" (2004), Giornate di studio, Cattolica, Novembre.
European University Institute, Robert Schuman Centre for Advanced Studies, Florence School of Regulation (2008), "Affordability in Basic Public Utilities: Regulation and Poverty Policies", Workshop, October the 3rd.
Ferrari, G. (1999), Household Welfare Comparisons by Area in Italy, *Bulletin of the International Statistical Institute*, 52nd Session, Tome LVIII, Book 1, Helsinki, p. 337-338.
Ferrari, G. (2000), Equivalent Incomes Based Upon Equivalence Scales as a Tool for Economic and Social Policy in Tuscany, Italy, *Proceedings of the 22nd SCORUS Conference on Regional and Urban Statistics and Research*, Shenzhen, China, 7-10 November.
Ferrari, G. (2003), Using Equivalence Scales as Spatial Deflators: Evidence in Inter-Household Welfare Regional Comparisons from Italian HBS Micro-Data, in *Household Behaviour, Equivalence Scales, Welfare and Poverty*, C. Dagum and G. Ferrari editors, Physica-Verlag, p. 273-294.
Ferrari, G., Maltagliati, M. (1999), Estimating and Comparing Equivalence Scales from Pooled Cross Section-Time Series of Expenditure Micro-data from the Italian Family Budget Survey 1987-1995, *Working Papers* N. 81, Dipartimento di Statistica, Università di Firenze, p. 1-32.
Ferrari, G., Maltagliati, M. (2000a), Confronti di benessere delle famiglie italiane, *Rivista di Economia, Demografia e Statistica*, Vol. LIV – N.1, gennaio-marzo, p. 87-100.
Ferrari, G., Maltagliati, M. (2000b), Estimating the Cost of Children for Divorced Parents in Tuscany, *Quantitative Methods in Economics (Multiple Criteria Decision Making X)*, The Slovak Society for Operations Research & University of Economics in Bratislava, November-December, p. 29-35.
ISAE (Istituto di Studi e Analisi Economica) (2007), Nota mensile, luglio, Roma.
Leser (1963), Forms of Engel Functions, *Econometrica*, 31 (4), 694-703.
Binh, Tran Nam & Whiteford, Peter, 1990. "Household Equivalence Scales: New Australian Estimates from the 1984 Household Expenditure Survey," The Economic Record, The Economic Society of Australia, vol. 66(194), pages 221-34, September.
Working, H., (1943), "Statistical Laws of Family Expenditure*", Journal of the American Statistical Association*, vol. 38, 43-56.

SUMMARY

The issues of household and individual well-being, poverty have in recent years attained a crucial relevance in the framework of globalization which has introduced problems that have never before arisen and proposed researches for explanations and solutions, which unfortunately are still far from being found, at least on an acceptable scale. In this paper, starting from "Banca d'Italia" budget survey data, we used two different approaches to estimate equivalence scales, in order to calculate equivalent incomes for each household of the sample. These equivalent incomes are then used to compare household welfare under territorial divisions.

Guido FERRARI, Dipartimento di Statistica Università di Firenze.
Mauro MALTAGLIATI, Dipartimento di Statistica Università di Firenze.

Rivista Italiana di Economia Demografia e Statistica Volume LXIII nn. 1-2 – Gennaio-Giugno 2009

MEASURING THE PROGRESS OF SOCIETIES

Enrico Giovannini, Adolfo Morrone, Giulia Ranuzzi

1. Introduction

Progress takes a variety of directions and forms. Therefore, throughout history, various notions of progress have gained favour depending on prevailing political regimes, cultural influences and environmental conditions. In the 18th century, morality, inventions and discoveries dominated people's views on progress. In the 19th century, spirituality took over. In the 20th century, progress was often equated with economic growth.

After the Great Depression and World War II, national accounting, and in particular Gross Domestic Product came to be seen by many as the main way of measuring progress. Although several alternative measures of well-being and societal progress have been developed by researchers during the '70s and the '80s (for example, the ones grouped under the so-called "social indicators movement"), it is only in the 1990s that initiatives concerned with sustainable development and measuring human development such as the UNDP Millennium Development Goals have captured the attention of media and have played a role in political debates. More recently, thanks to initiatives carried out by some national and local political authorities, the research on the measurement of quality-of-life and happiness, and initiatives undertaken by the OECD, as well as other international organisations, a new movement aiming at measuring progress is emerging.

This paper describes, first of all, the main characteristics of this new movement towards the measurement of the progress of our societies. Second, it proposes a possible framework for the measurement of societal progress in terms of "sustainable and equitable well-being", where economic, social and environmental dimensions are integrated with other aspects of human well-being. Finally, the paper describes the main characteristics of "WikiProgress", a forthcoming internet platform under development in the context of the Global Project on "Measuring the Progress of Societies" launched in 2008 and hosted by the OECD.

2. The Istanbul Declaration

In June 2007, three years after its 1st World Forum on "Statistics, Knowledge and Policy" held in Italy, the OECD, in collaboration with other international organisations, ran the 2nd World Forum on "Measuring and Fostering the Progress of Societies". Some 1200 people, from over 130 countries attended. Presidents and ministers mixed with civil society leaders, captains of industry met the heads of charitable foundations and leading academics. They all shared a common interest in wanting to develop better measures of how the world is progressing.

The conference led to the "Istanbul Declaration", signed by the European Commission, the Organisation of the Islamic Countries, the OECD, the United Nations, the Organisation of the Islamic Conference, the United Nations Development Programme, UNICEF, UNESCO, the United Nations Fund for Partnership, the World Bank, and several other organisations[1].

The Declaration calls for action to identify what "progress" means in the 21st century and to stimulate international debate, based on solid statistical data and indicators on both global issues of societal progress and how societies compare. In a nutshell, the Declaration calls for actions to:

- Encourage communities to consider for themselves what "progress" means in the 21st century;
- Share best practices on the measurement of societal progress and increase the awareness of the need to do so using sound and reliable methodologies;
- Stimulate international debate, based on solid statistical data and indicators, on both global issues of societal progress and comparisons of such progress;
- Produce a broader, shared, public understanding of changing conditions, while highlighting areas of significant change or inadequate knowledge;
- Advocate appropriate investment in building statistical capacity, especially in developing countries, to improve the availability of data and indicators needed to guide development programs and report on progress toward international goals, such as the Millennium Development Goals.

The Istanbul Declaration marks an important moment in the history of a movement to go "Beyond GDP" that grew dramatically over the last decade. As documented by the proceedings of the 2004 and 2007 OECD World Forum, as well

[1] See http://www.oecd.org/dataoecd/23/54/39558011.pdf

as by several presentations given at the preparatory conferences organised towards these two events and the "Knowledge Base" available at www.oecd.org/progress, the number of initiatives launched in this decade to measure progress of countries and local communities is astonishing and growing every year. The US-based "Community Indicators Consortium", the French "Forum for Other Indicators of Wealth (FAIR)", the Latin American initiative "Como Vamos", the Italian network "Sbilanciamoci", the UK initiative on the measurement of wellbeing of local communities, the reports promoted by the Australian and Irish statistical offices on measuring the progress of their societies, the "State of USA" and the "Canadian Index of Well-Being" initiatives are just few examples of a growing movement.

Most of these initiatives share the following common features:

- Their aim is to give, through statistical indicators, a holistic view of how a country or a local community is performing, beyond the classical economic results;
- The measurement framework they use is built through a wide consultation with stakeholders. Although this approach does not have necessarily the "theoretical rigour" of an academic-led research, it may provide a good view of what people values most, with a significant increase in the legitimacy of the initiative vis-à-vis citizens and policy makers;
- In some cases, especially for initiatives that focus on the local dimension, the measurement of societal progress action-oriented, aimed at building policies able to address the concerns raised by the indicator sets and at increasing the accountability of public institutions. These community-led initiatives link the measurement dimension with the political one. This characteristic is important as it allows filling an important gap of previous academic-led initiatives.
- Beside their differences in the measurement frameworks, these initiatives can be seen together as a new and powerful way to improve the functioning of a democracy in the "information age".

3. The Global Project on "Measuring the Progress of Societies" and its impact on the international political debate

The *Global Project on Measuring the Progress of Societies* is an OECD-hosted project that for the past few years has been looking at new sets of economic, social and environmental indicators with the aim to provide a comprehensive picture of how our world is really performing and of how the well-being of a society is evolving.

The Global Project is an international network of organisations from all sectors of society with partners including the World Bank, United Nations Development Programme, UNICEF, international development banks, the European Commission and associations with research institutes, Non-Governmental Organisations (NGOs), and statistical offices from both developing and developed countries. The Project has three main goals:

- *What to measure?* In order to measure progress it is essential to know what it looks like. For this reason, the project encourages debate about what progress means in different societies. The Project is developing methods and guidelines to carry out these debates effectively.
- *How to measure progress?* The Project is developing best practices in how to measure progress and its component parts, some of which are not yet measured well using existing statistical indicators.
- *Ensuring new progress measures are used*. New ICT tools offer a great potential to turn information into knowledge and to disseminate this knowledge among a much broader pool of citizens than those who currently have access to such information. The Project is developing new tools for public use.

Current streams of work include formulating guidelines for the development of "measuring progress" initiatives and measurements not yet covered by international statistical standards; designing communication tools to increase the accessibility of progress indicator data; establishing dialogues with governments; civil society and statistical offices on progress measurement; and identifying lessons for successful sets of indicators.

The Project has been able to build a real community on the issue of measuring societal progress, well-being, sustainable development, quality of life, etc, while capturing the attention of policy makers and political leaders. For example, in 2008 the French President Nicolas Sarkozy established a Commission on the "Measurement of Economic Performance and Social Progress". Led by Prof. J. Stiglitz and participated in by four other Nobel Laureates and well-known experts from all over the world (http://www.stiglitz-sen-fitoussi.fr)[2], the Commission is addressing the limitations of current statistical frameworks (for example, national accounts) to provide meaningful measures of societal well-being in the short and long term, and is developing research work to overcome such limitations. The Commission will present its final report in September 2009.

The main concern that led President Sarkozy to establish the Commission was the growing distance between current measures of economic performance, in

[2] Chaired by Stiglitz and with Amartya Sen as Chair Adviser and Jean-Paul Fitoussi as Coordinator.

particular those based on GDP figures, and people's perceptions about the quality of their life. "*This gap is so large and so universal that it cannot be explained by reference to money illusion and/or to psychological characteristics of human nature. The issue here is both analytical and political, and current statistical systems, which may have served us well in a not too distant past, are in need of serious revisions*"[3]. To organise its work, the Commission selected three main directions of study which correspond to three of the already identified main causes of divergences between perceptions and measures:

- Classical GDP issues: limits of GDP as an indicator of socio-economic progress or economic performance can be addressed by investigating possible extensions or modifications of the current conceptual framework;
- Quality of life: this direction of study covers the measurement of social progress taking into account broader perspectives on well-being, including metrics derived from asking people how they themselves feel;
- Sustainable development and environment: one of the biggest concerns about current measures of economic performance and social progress is related to sustainability and one of the areas where sustainability is most questioned is the environment.

An important outcome of the Commission's work will be several suggestions for alternative indicators which may provide a better description of economic performance and social progress. Taking stock of similar work conducted in the past, the Commission will be cautious about the number of indicators proposed. Here, as elsewhere in economics, there are trade-offs: a larger number of indicators may better reflect the diversity of issues and individual situations, but an excessively large number may provide a confused picture of the overall situation. On the other hand, a single figure mixing a large number of socio-economic phenomena provides an inadequate basis for appropriate policy measures.

[3] See the "Issues paper" available on the Commission's web site www.stiglitz-sen-fitoussi.fr.

In June 2008, the World Economic Forum established a Global Council on "Benchmarking the progress in societies", with the participation of experts from several institutions[4]. The OECD Chief Statistician is the chair of the Council. Four key issues are being considered by the Council:

- which metrics are used by the different benchmarking tools and which are the indicators and indices most widely used as monitoring and policy instruments at the national level?
- which are the variables that could complement measures of income and build a more nuanced and accurate understanding of economic and societal progress? Can a common definition of well-being, and a common set of indicators, be agreed upon?
- what can be done to improve statistical capability, comparability and reliability of data across countries?
- how can data be made available to the general public more effectively in order to promote openness and government accountability? Can the foundations and organisations established to promote statistics literacy and dissemination to the civil society in a number of countries be replicated elsewhere? How can their actions be made more effective?

4. A framework to measure the progress of societies

Over the last three decades, several frameworks have been developed to measure well-being, quality of life, societal progress and development. Some of these frameworks use a conceptual approach and are derived from a particular view of what progress means, while others use a political approach in which the components of progress are selected through political consultations and agreements.

However, although we should expect and value different views of progress, the lack of a common starting point delays research and often leads to duplication of efforts. Therefore, the Global Project is proposing a comprehensive framework, which does not wish to create a single view of what progress is, but hopes to

[4] N. Burnett (UNESCO), P. Cheung (United Nations), D. C. Esty (Yale University), R. N. Garcia (IMCO), E. Giovannini (OECD), D. Kauffman (World Bank Institute), J. A. Kurtzam (Kurtzam Group), R. A. Lawson (Capital University), J. Olaya (Transparency International), R. Layard (London School of Economics), H. Rosling (Karolinska Institutet), A. M. Said Aly (Al-Ahram Centre), Ruut Veenhoven (Erasmus University Rotterdam), S. Young (International Labour Office), D. Farrell (McKinsey & Company), Kemal Dervis (UNDP) and Ian Ayres (Yale Law School).

provide a starting point to facilitate future research for initiatives that aim to measure progress at local or national level.

4.1 *Can we reach agreement on a framework?*

Frameworks are a tool to focus and clarify the scope of an enquiry. They facilitate this by delineating the dimensions used to build up a particular concept and creating a logical structure that illustrates how these dimensions relate to one another. A framework for the measurement of progress should underpin the public discourse about the level of a community's well-being; orient political choices and make them more evidence-based; it should also make policymakers accountable to citizens.

As already noted, measurement frameworks can be developed following two approaches: in the first case, they are simply developed through political negotiation and tend to have a pyramidal structure based on headline, sectorial and detailed indicators (e.g. frameworks for sustainable development) or else on goals and target indicators (e.g. the Millennium Development Goals). The second approach looks at relevant scientific literature, trying to define some broad domains of progress and then divide them into potential dimensions. These dimensions are more specific building blocks that should correspond to what people value most according to empirical surveys.

Of course, diversity of core values processes and languages have led to the development and application of different frameworks for societal progress. These differences range from the conceptualisation of what progress is to the choice of dimensions to include, to the ways in which dimensions are defined, to the relationships between dimensions and to the weight to be attributed to each dimension (or whether dimensions of progress should be weighted at all).

Horace argued that "there might be as many preferences as there are people". With that in mind, one may doubt the very possibility of rational collective choice. But as Sen (1999) suggests, if we have enough information and do not aim for too much precision, we can take into account the diversity of the preferences, interests, concerns and predicaments of different members of society in order to produce a reasoned and democratic social choice. Such a framework should be broad enough to allow its users to define progress from their own perspectives based on their value systems and what they view as relevant to their lives and those of their family, friends or community. While it is important to maintain relevance, to achieve such a level of generality a framework requires substantial abstraction and simplification.

In conclusion, a framework that aims at assessing the progress of societies should have the following characteristics:

- It should be built on solid conceptual ground;
- It should contain broad domains and potential dimensions that must be incommensurable, irreducible, non-hierarchical and valuable;
- It should not require too much precision (Sen, 1999), nor should it be too prescriptive;
- It should focus on outcomes (or ends) rather than outputs (or means);
- The process of its development should involve public participation and dialog among relevant stakeholders for greater legitimacy.

4.2 *A proposed framework to measure the progress of societies: domains and dimensions*

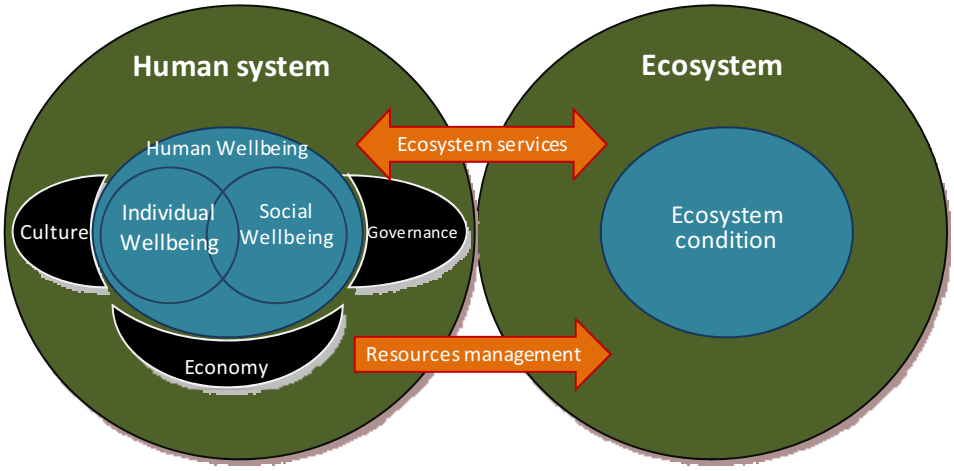
The proposed framework aims to select and present the key measures of societal progress, not to construct a model of how the world works. Many aspects of life affect societal progress and individual well-being and this framework does not seek to account for all of them. Instead, we seek to select a set of dimensions of societal progress that can be influenced by human beings. Earthquakes, for example, have an impact on both people and the environment, but societies cannot influence the number and intensity of earthquakes, even if they can provide safer houses in earthquake zones (or not build there in the first place). On the contrary societies should reduce people's vulnerability to earthquakes and measures of societal progress should pick up this effort rather than the number of earthquakes per se.

As we are seeking a framework that is both broad and flexible, something that will provide a solid foundation for others to adapt to their own purposes, several leading frameworks have been analysed to formulate the framework in figure 1. This framework considers that societies are based on two systems: the Human system and the Ecosystem[5]. They are linked through two different channels, "Resource management" and "Ecosystem services". Resources management represents the effects of the human system on the ecosystem, through resource depletion and pollution. Ecosystem services link the two systems in both directions. The ecosystem benefits the human system through positives services like food, clean water. But it can also do damage through things earthquakes and

[5] The proposed framework largely draws from the model presented by Robert Prescott-Allen at a Conference on "Measuring Wellbeing and Societal Progress" organised in 2006 by the OECD, the Joint Research Centre of the European Commission and the Centre for Research on Lifelong Learning.

floods. The human system may also provide positive services to the ecosystem (or its capacity for supporting life) through providing food and water for wild animals in times of hardship, tackling invasive species and so on.

Fig. 1 – *The proposed framework of the progress of societies.*



Human wellbeing is the key domain and its dimensions represent Alkire's "reasons for action": therefore, in our framework it comprises the core human ends that societies pursue. An increase in human wellbeing is the final goal of progress.

Human wellbeing can be considered as comprising individual and social outcomes. In fact, human wellbeing may be conceived as a collection of attributes that characterise the kind of life that each person pursues, and their level of freedom (with 'freedom' used in the sense of Sen who takes it to be the range of opportunities open to people). Some of these attributes will be specific to each person (one's own state of health, knowledge, etc.) and can be clustered together as attributes of "individual wellbeing". Other attributes are shared with other people (those living within the same family or neighbourhood), or reflect the relations between them (e.g. the extent and quality of relationships with others), or how a society is peaceful, resilient, cohesive, and can be clustered together as "social wellbeing".

Human wellbeing is supported by three domains: economy, culture and governance. These are seen as important insofar as they are key supporting pillars to human wellbeing, rather than seen important for their own sake. Having a strong economy, effective governance and vibrant culture is not wellbeing in itself, but

these factors do – typically – provide an enabling environment in which human wellbeing will improve. Therefore, they are considered "intermediate goals".

The ecosystem has only one domain (ecosystem condition), which represent the wellbeing of the ecosystem. Ecosystem wellbeing is equally important if one sees the ecosystem as important in its own right or if one takes a more anthropocentric view (where one sees the ecosystem as important simply because it provides the human system with resources and services which contribute to human wellbeing).

At this point, one could define:

- the "wellbeing of a society" (or societal wellbeing) as the sum of the human wellbeing and the ecosystem condition; and
- "progress of a society" (or societal progress) as the improvement in human wellbeing.

But it is also important to recognise the role played by inequalities in human wellbeing and ecosystem condition across and within societies or geographical regions and between generations. Consider, for instance, an average increase of the material wellbeing of a society, but an increase which goes solely to the richest 10% of people, while the material wellbeing of the poorest 10% declines. The average level of material wellbeing may have risen, but has their really been progress in the society? Similar arguments can be applied to the sustainability dimension, i.e. the distribution of wellbeing between generations. Therefore, we believe that the wellbeing of a society also depends on the way in which the various items that shape people's lives are distributed in society and it cannot be assessed without considering its sustainability over time and/or the wellbeing of the future generations.

Putting the first two and second two pairs of points together we define societal progress as occurring when there is an improvement in the "sustainable and equitable wellbeing of a society".

The framework we propose does not simply equate progress to an increase in individuals' evaluations of happiness/life satisfaction, though it sees this as an important element, but also underlines the importance of objective conditions and economic, social and environmental achievements (see box below). It puts emphasis on the importance of the wellbeing of the current generation, but also defines progress as an increase in equitable and sustainable wellbeing, thereby recognising that not all individuals are properly equipped or informed to take a long-term perspective. Finally, it looks compatible with Sen's capabilities approach stressing the fact that to enhance human wellbeing the intermediate goals of the human system (economy, governance and culture) should provide conditions under which individuals can make use of their potentials. The proposed framework could be, thus, defined as "eclectic".

Of course, a framework based on these broad domains of progress is not immediately operational. To become useful for those who want to measure societal progress, more precise dimensions need to be defined, within the domains contained in Fig. 1. To define these dimensions we have looked at work from around the world, much of which is contained in the Knowledge Base available on www.oecd.org/progress. The result of this analysis led us to a set of "final goals for progress" (covering human wellbeing and ecosystem condition) and a set of "intermediate goals" (covering economy, culture and governance). The "final goals" are direct measures of human and environmental wellbeing, while the "intermediate goals" are those elements that are key inputs into human and environmental wellbeing. Finally, the links between the two sets of goals need to be considered, as well as two key "cross-cutting perspectives", i.e. the intra-generational (equity) and the inter-generational (sustainability) perspectives.

The importance of inequality for people's wellbeing is widely recognised by theories of welfare, empirical research and social norms. For example, recent literature on subjective wellbeing has found some correlation between life satisfaction (i.e. subjective wellbeing) and income inequality and insecurity. But these areas are not included as separate "dimensions" of social progress, because they cut across multiple dimensions of progress[6]. We recognise that an equitable (however defined) distribution of resources and social outcomes between individuals, population groups (for instance, gender equality) and between generations is an important societal goal[7]. Such considerations need to happen

[6] Poverty, for example, can be defined in several ways: one could use the term to mean a lack of financial resources; some consider that it crosses multiple dimensions (Sen, 1992). According to Marco Mira D'Ercole (2009, forthcoming), "poverty is a complex phenomenon, varying across time and space, with different philosophical perspectives leading to different conclusions about its nature, and with alternative measures sometimes providing conflicting indications about its size and evolution". Therefore, some might turn to indicators of income and wealth distribution to measure poverty. Others might look more broadly. But it is the basket of these measures that would be used to assess multidimensional poverty.

[7] As "progress" is a dynamic concept, its measurement can be made looking at the temporal movements of aggregate indicators, based on averages or other summary measures. However, as questions about people's wellbeing are ultimately about the lives of individuals in society, we cannot really evaluate a distributional change without knowing, for example, if the formerly underprivileged remain at the bottom of the heap or have exchanged places with the more privileged. While it is unrealistic to obtain indices of every individual's views about progress, it could be useful to conduct longitudinal studies to understand whether and under what circumstances people experience different outcomes—do the same individuals remain poor over the years, or is poverty transient?

throughout the framework, and can be applied to analyse the conditions of specific population groups (children, women, etc.).

1. Similar considerations apply to sustainability/vulnerability/resilience. A person can be healthy, educated and have a job today, but be at risk tomorrow because of poor investment decisions, or a weak social security system. Vulnerability has an inter-temporal dimension and can be important for individuals, as well as specific social groups (e.g. farmers who live in regions subject to drought). Such inter-temporal considerations can be taken throughout the framework.

FINAL GOALS

Ecosystem Condition: outcomes for the environment

- land (geosphere)
- freshwater, oceans and seas (hydrosphere)
- biodiversity (biosphere)
- air (atmosphere)

Human wellbeing: outcomes for people

- physical and mental health
- knowledge and understanding
- work[8]
- material wellbeing
- freedom and self-determination
- interpersonal relationships

INTERMEDIATE GOALS

Economy

- national income
- national wealth

[8] This dimension should take into account not only the availability of work but also work conditions and the availability of decent work according to the definition given by ILO.

Governance

- human rights
- civic and political engagement
- security and violence
- trust
- access to services

Culture

- cultural heritage
- arts and leisure

LINKS BETWEEN THE TWO SETS OF GOALS

Resource management, use, development and protection

- resource extraction and consumption
- pollution
- protection and conservation of economic and environmental assets

Ecosystem services

- resources and processes provided
- impact of natural events

CROSS-CUTTING PERSPECTIVES

Intra-generational aspects: equity/inequality

Inter-generational aspects: sustainability/vulnerability/resilience

5. Towards WikiProgress, a global platform to measure and assess societal progress

As already mentioned, a huge number of initiatives aimed at measuring economic, social and environmental developments, through statistical indicators, have been identified. Each of these initiatives uses its own "taxonomy" of progress (quality of life, sustainable development, etc.), as well as its own set of statistical

measures. Although a complete analysis of these taxonomies has not yet been carried out, the impression is that in many cases the main dimensions of progress considered are not so different.

While the Global Project aims at creating national roundtables to identify what progress means for a given country, it cannot expand its activity to reach all sub-national initiatives, which, in turn, often provide very interesting ideas and practices on how to build, disseminate and use indicators. Therefore, one of the key questions for the future work of the Project is how to link the sub-national level with the national and supranational levels. Another key challenge of the Global Project is to stimulate, on a worldwide level, an exchange of best practices on how to establish progress initiatives, to build and disseminate indicators and to engage stakeholders in such activities. Finally, given the global dimension of the Project, the "dream" of building a single repository of progress indicators, where users can compare situations for two or more countries, regions or local communities, should be accomplished.

To address all these issues, the Global Project is working towards the establishment of a global platform to serve all people in the world to understand and debate, using statistical indicators, whether the world itself, or a particular country or region, is making progress. Given all the objectives mentioned above, as well as considering the technical and resource constraints, such a platform cannot follow the "classical" approach of web building. Fortunately, the development of Web 2.0 tools makes the problems less insurmountable. Thanks to the contact established in 2007 with several very innovative companies and government agencies active in the field of ICT, the idea of building a wiki platform, "WikiProgress", emerged in June 2007 during the Istanbul Forum.

The main attribute of a Wiki is that authors contribute their knowledge to a single repository, designed to represent the synthesis of what the "collective intelligence" is able to build about that particular subject. As reported on Wikipedia, a wiki:

- "invites all users to edit any page or to create new pages within the wiki Web site, using only a Web browser without any extra add-ons;
- promotes meaningful topic associations between different pages by making page link creation almost intuitively easy and showing whether an intended target page exists or not;
- seeks to involve the visitor in an ongoing process of creation and collaboration that constantly changes the Web site landscape".

A key characteristic of wiki technology is the ease with which pages can be created and updated. Generally, there is no review before modifications are accepted. Many wikis are open to alteration by the general public without requiring them to register, while private wikis require user authentication to edit pages, and

sometimes even to read them. WikiProgress should bring data and metadata in the same environment to build a single, integrated database. Moreover, it should allow a simple use of statistical data to create charts and tables to be posted on its pages. The development of a relational database whose dimensions evolve over time following a "bottom-up" process poses immense difficulties from a technical point of view, but the Project has identified a possible solution and is working towards the development of a first pilot, to launch the platform in 2009.

As already mentioned, WikiProgress should be able to represent the catalyst of initiatives existing around the world on the measurement of progress, as well as their use for raising awareness amongst stakeholders, informing them on key economic, social and environmental trends and allowing them to discuss relevant issues based on solid evidence. Therefore, while Wikipedia answers questions like "Who is this person?", "What is this?" and so on, WikiProgress should mainly answer the following questions:

- Who is developing initiatives on measuring progress (well-being, quality of life, etc.);
- What type of framework do these initiatives use?
- Which indicators are being used to measure the different dimensions of progress?
- How is my country/region/community achieving over time and in comparison to other similar territories?

Moreover, WikiProgress should represent "the" place where both experts and practitioners could share their practices on indicator design, calculation and dissemination, as well as where stakeholders interested in developing initiatives in this field can find reference documents and assistance on how to establish progress initiatives, design websites, download software, etc. Finally, WikiProgress should represent a tool for the Global Project, to enable extraction from existing initiatives of the information necessary to identify good practices, run comparative studies, show similarities and compare the differences between various initiatives, etc.

To reach these objectives, WikiProgress should be a multi-purpose website, with two main parts:

1. A "Classical Wiki" where users can find:
    - materials developed by the Project (proceedings of conferences, handbooks and guidelines, software, etc.), contributed by the Partners and the Associates to the Global Project;
    - information about existing or new initiatives aimed at measuring progress around the world, contributed by those who run these initiatives and/or other people.

2. A "Statistical Wiki", where data and metadata can be shared, assessed and eventually uploaded at the end of a "quality-assurance" process. In particular, the user should be able to:
   - upload data and metadata and submit them for quality evaluation;
   - navigate the database by country (map and list) or by topic (taxonomy and list), exploring data and metadata;
   - create tables and charts and export them in various formats to populate texts, blogs, wikis, etc.

Work is underway to develop the first prototype on time for the Third OECD World Forum on "Statistics, Knowledge and Policy", to be held in Busan (Korea) on 27-30 October 2009. The prototype will assemble components which already exist, developed by the UNICEF, the OECD and the NCVA, a Swedish-based research centre. It will also build on the experience the OECD has gained with "Wikigender" (www.wikigender/org), the platform launched in 2008 by the OECD Development Centre to foster the analysis of gender disparities using both textual and statistical information.

6. Conclusion

This paper highlighted some of the activities carried out in the context of the Global Project on "Measuring the progress of societies" officially launched by the OECD in 2007. In particular, it presented the main characteristics of the Taxonomy of Societal Progress currently under development, which defines progress as "an improvement in the sustainable and equitable well-being of a society". The proposal encompasses several other frameworks developed over the years by researchers and practitioners around the world.

However, the proposed Taxonomy is not a prescriptive one. Recognising that the concept of progress may differ depending on different historical and cultural conditions, its aim is to help those who want to measure progress in practice to avoid re-inventing the wheel, but rather to adapt this framework to their own needs.

To be valuable, the Taxonomy needs to evolve, incorporating the results of the international community of researchers and practitioners. WikiProgress has been designed to be enriched by the contributions of this community. Their contributions will be used to further refine the Taxonomy and its components, as well as to evaluate new proposals to measure progress in all its dimensions.

References

ABS, (2004), "Measures of Australia's Progress", *The Australian Bureau of Statistics*, Canberra.
Alkire, S., (2002), "Dimensions of Human Development", *World Development,* 30(2), pp. 181-205.
Alkire, S., (2008), "The Capability Approach: Mapping Measurement Issues and Choosing Dimensions", in Kakwani N. and Silber J., (eds.), *The Many Dimensions of Poverty*, Palgrave MacMillan, New York.
Becker, G. S., (1981), "Altruism in the Family and Selfishness in the Market Place", *Economica*, 48(189), pp. 1-15.
Berger-Schmitt, R. and Noll, H.H. (2000), "Conceptual Framework and Structure of a European System of Social Indicators", *EU Reporting Working Paper,* N. 14, Centre for Survey Research and Methodology, Mannheim.
Blinder, A. S., Krueger A. B., (2004), "What does the public know about economic policy and how does it know it?", *Working Paper 10787,* National Bureau of Economic Research.
Brekke, K. A., Lurås, H. and Nyborg, K., (1996), "Allowing Disagreement in Evaluations of Social Welfare" *Journal of Economics*, 63(3), pp. 303-324.
Chambers, S. and Jopstein, J., (2001), "Bad Civil Society", *Political Theory*, 29(6), pp. 837-865.
Confucius, (ca.500BC), *The Great Learning*, translated and reprinted by Kessinger Publishing, Llc, Whitefish, MT.
Costanza, R., et al., (2009), "Beyond GDP: The Need for New Measures of Progress", *The Pardee Papers,* N. 4, Boston University.
Cox, J. C., (1997), "On testing the Utility Hypothesis", *The Economic Journal*, 107(443), pp. 1054-1078.
De Spinoza, B., (1676), *A Theologico-Political Treatise,* translated by Elwes, R., Dodo Press, London.
De Vries, W. F. M., (2001), "Meaningful Measures: Indicators on Progress, Progress on Indicators", *International Statistical Review*, 69(2), pp. 313-331.
Dyson, M., (1973), "Odes iii.14", *Greece and Rome*, 20(2), pp. 169-179.
Einhorn, H. J. and Hogarth, R.M., (1980), "Backward and forward thinking in decision making", *Selected paper 62*, Center for Decision Research, University of Chicago, Chicago.
European Commission, (2007), "Measuring progress towards a more sustainable Europe", *The EU sustainable development strategy,* Eurostat Statistical Books, Luxembourg.
Eurostat, (2000), "Cultural statistics in the EU. Final report of the LEG*", Eurostat Working Paper*, No. 1., Eurostat, Luxembourg.
Fogel, R. W., (2004), *The Escape from hunger and Premature Death 1700-2100,* Cambridge University Press, Cambridge.
Fogel, R. W., (2005), "Reconsidering Expectations of Economic Growth after World War II from the Perspective of 2004", *Working Paper 11125*, National Bureau of Economic Research.
Fogel, R.W., (1999), "Catching up with the Economy", *The American Economic Review*, 89(1), pp. 1-21.
Garonna, P., and Triacca, U., (1999), "Social Change: Measurement and Theory", *International Statistical Review*, 67(1), pp. 49-62.
Giovannini, E., (2004), "Accounting Frameworks for Sustainable Development: What Have We Learnt?", *Measuring Sustainable Development: Integrated Economic, Environmental and Social Frameworks,* OECD, Paris.

Goebel, B. L., Brown D. R., (1981), "Age differences in motivation related to Maslow's need hierarchy", *Developmental Psychology,* 17, 809-815.
Grange, J., (1996), "The Disappearance of the Public Good: Confucius, Dewey, Rorty", *Philosophy East and West*, 46(3), pp. 351-366.
Hall, J., (2005), "Measuring Progress: An Australian Travelogue", *Journal of Official Statistics*, 21(4), pp. 727-746.
Itay, A., (2009), "Conceptions of Progress: How is Progress Perceived? Mainstream Versus Alternative Conceptions of Progress", *Social Indicators Research*, 92(3), pp. 529-550.
Juster, F. T., 1985. Preferences for work and leisure. In: F.T. Juster and F.P. Stafford, eds. 1985. *Time, Goods, and Well-Being.* Ann Arbor: University of Michigan*,* Institute for Social Research.
Katzner, D. W., (1979), *Choice and the Quality of Life,* Sage Publications, Beverly Hill.
Landefeld, S., 2000. *GDP: One of the Great Inventions of the 20th Century.* United States Department of Commerce: Bureau of Economic Analysis.
Meda, D., (2009), *What Progress Should Be Measured*, OECD, Paris.
Michalos, A. C., (1997), "Combining social, economic and environmental indicators to measure sustainable human well-being", *Social Indicators Research,* 40, 221–258.
Michalos, A., et al., (2008), "*An Approach to the Canadian Index of Wellbeing*", Atkinson Charitable Foundation.
Mira D'ercole, M., (in press). "*Alternative Measures of Poverty",* Extract from a forthcoming *Handbook on Measuring Progress in Practice*, OECD, Paris. (Due for publication in 2009).
Naussbaum, M. and Sen, A., (1993), *The quality of life,* Oxford University Press, Oxford.
Osberg, L. and Sharpe, A., (2001), "Trends in Economic Well-being in Canada in the 1990s", *The Review of Economic Performance and Social Progress*, Institute for Research on Public Policy and the Centre for the Study of Living Standards, pp. 233-246.
Ottaviani, G. M., (2002), "Statistics, from a Tool for State and Society to a Tool for All Citizens", *International Statistical Review*, 70(1), pp. 30-32.
Pettijohn, T. F. II, and Pettijohn, T.F., (1996), "Perceived Happiness of College Students Measured by Maslow's Hierarchy of Needs", *Psychological Reports*, 79, 759-62.
Prescott-Allen, R., (2006), "The structure of a well-being index", *JRC/OECD workshop*, Milano.
Schumacher, E.F., (1973), *Small is Beautiful: Economics as if People Mattered*, Harper and Row, New York.
Scoble, M. J., (2004), "Unitary or Unified Taxonomy?", *Philosophical Transactions: Biological Sciences*, 359(1444), pp. 699-710.
Sen, A., (1992), *Inequality re-examined*, Clarendon Press, Oxford.
Sen, A., (1999), "The Possibility of Social Choice", *The American Economic Review*, 89(3), pp. 349-378.
Sen, A., (1999), *Development as Freedom,* University Press, Oxford.
Stevens, C., (2005), "Measuring Sustainable Development", *OECD Statistics Brief*, No. 10, OECD, Paris.
UK Government, (1999), *Quality of Life Counts,* Online Publication from the Sustainable Development Unit.
Unesco Institute for Statistics, (2009), *The 2009 UNESCO framework for cultural statistics,* UNESCO, Paris.

SUMMARY

In June 2007, the OECD, in collaboration with other international organisations, ran the 2nd World Forum in Istanbul on "Measuring and Fostering the Progress of Societies". This conference led to the Istanbul Declaration, signed by several international organisations who affirmed their "commitment to measuring and fostering the progress of societies" in all its dimensions and encouraged the OECD to begin a Global Project on "Measuring the progress of societies" (see www.oecd.org/progress).

Since then, numerous initiatives have been launched around the world to measure progress/well-being/sustainable development/happiness of countries and local communities. The research work has made substantial progress, especially in areas not well covered by international statistical standards. The paper, showing the streamlines of work of the Global Project, will review the state of the art in this domain identifying whether any consensus on how to measure well-being and societal progress is emerging.

Enrico GIOVANNINI, Organisation for Economic Co-operation and Development (OECD).
Adolfo MORRONE , Organisation for Economic Co-operation and Development (OECD).
Giulia RANUZZI, Organisation for Economic Co-operation and Development (OECD).

Rivista Italiana di Economia Demografia e Statistica Volume LXIII nn. 1-2 – Gennaio-Giugno 2009

POVERTY AND SOCIAL EXCLUSION IN SPAIN: REGIONS AND RISK GROUPS

Vicente Gozálvez Pérez, José Ramón Valero Escandell[1]

1. Introduction

Spain is currently enjoying a period of comparatively strong growth in global terms, with little exposure to extreme poverty. However, the country still has marked social inequalities, and social exclusion is becoming an increasingly complex issue. All the poverty indicators that are published are at a nationwide or excessively wide regional scale, when the greatest inequalities in living conditions are found on a much smaller scale and affect specific social collectives. The differences between areas in any one city are today much greater than those that exist between Spain's various autonomous regions.

2. Poverty and exclusion: differences and inter-relationships. Sources of information

Poverty and social exclusion are closely related: extreme, absolute or severe poverty is the simplest form of serious exclusion, as well as the most profound cause for many other forms of exclusion that would seem to have no apparent connection to poverty.

In the EU, the concept of poverty is applied to those whose resources are so limited that they do not cover the minimum standard of living considered acceptable by their fellow citizens; it affects people who endure situations of relative or moderate poverty who do not necessarily display behaviours of social exclusion, although they are at an increased risk of falling into this category, or into

[1] This research was carried out as part of the CSO2008-01796 project, titled "Family reunification among African and Latin-American immigrants in Mediterranean Spain", directed by Dr Vicente Gozálvez Pérez and funded by the Ministry of Science and Innovation, State Department for Research.
The maps were produced with the collaboration of Xavier Amat Montesinos and Gabino Martín-Serrano Rodríguez, from the Human Geography Department.

more serious levels of poverty.

Social exclusion means a "*social process that separates an individual or a group from job-related, economic, political or cultural possibilities which others have access to and enjoy*" (Giner, 2006, 324). It indicates limitations for some groups or individuals to participate in certain social circles; furthermore, and as stated by Subirats (2005, 176), its analysis requires new dynamic and multi-dimensional tools to observe different risk and vulnerability factors, by identifying collectives and quantifying the effect on them.

Whereas sociologists tend to be most interested in setting effective indicators in order to detect homes at risk of social exclusion, geographers are particularly concerned with questions of scale. Poverty indicators (whether suitable or limited) are usually produced on a national or regional scale, rather than a city-wide or intra-urban scale. As a result, a country with strong internal inequalities will show much greater differences than a homogenous country (even though this second country may be in a more precarious position), and a slight improvement or reduction in social provision (in the basic pension or minimum wage) produces notable variations in the percentage of people above a certain poverty threshold.

It is of little help when analysing groups and specific areas that the *Living Conditions Survey* (LCS), its predecessor the *European Community Household Panel* and the *Family Budget Survey*, among others, all give national or regional data, and fail to show the real situation of poverty and exclusion at a local (infra-municipal) level. Today, the internal differences in any one city are greater than those between provinces. The size of the statistical sample is acceptable at a state level, but limited and insufficient at smaller scales. For specific situations, field work must be relied upon, which is almost never co-ordinated or homogenous for the whole of the country.

The existing demographic data are also inadequate. The *Population and Household Census* provides highly valuable, detailed and accessible information at very different scales, but the latest is from 2001, and Spain has experienced a major transformation since then. The *Municipal Register of Inhabitants*, with data from January of each year, provides only a small number of variables. Data from the respective Education, Interior and Work ministries, which are more recent but lack direct information on poverty, provide some insight into collectives at risk, such as the elderly, foreign immigrants, the unemployed, prison inmates and school children.

Almost all the data on poverty and social exclusion provide figures for basic demographic variables (type of residence, gender, age, household size and structure), but not on ethnic groups or nationality, thus significantly reducing any analysis of the specific situation of gypsies or immigrants, for example. The Spanish authorities, and those of many other countries, do not readily admit to how

widespread certain problematic situations can be, due to the additional social demands that this could lead to (Alonso, 1995, 91).

In the midst of a crisis such as we face today (as of May 2009), any delay in the appearance of statistical data is felt all the more keenly. The most recent LCS was published in 2007 with data from 2006, and comparisons in the EU as a whole are even older. Job levels during the first years of the 21st century rose more in Spain than in any other EU country, and unemployment is now growing at a much greater rate than in the rest of Europe, affecting mainly the lower paid and immigrants. Logically, the consequences with regard to the degree of poverty and the risk of social exclusion are clear, but the data to back this up will not be available for some time. The analysis presented here should be understood within that context.

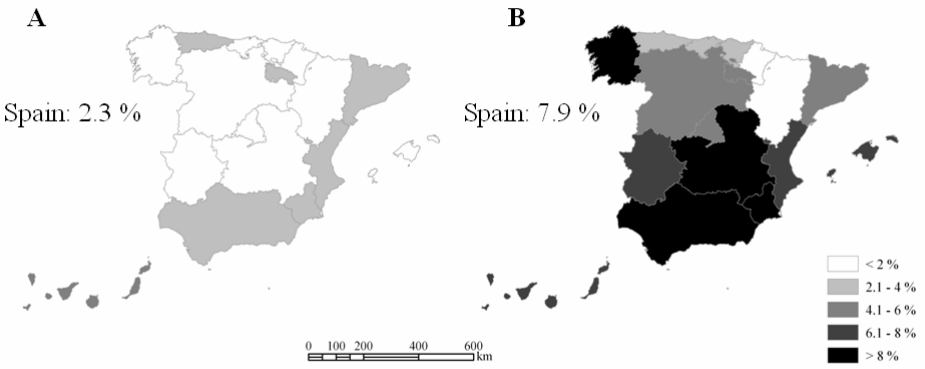
3. Reduction and recovery of poverty and severe exclusion

The distinction between absolute and relative poverty is essential to understand the case of Spain. Absolute poverty has been reduced drastically in recent decades, in a country that has experienced sharp increases in development in recent years: the latest UNDP *Human Development Report* (from 2007-08, but with data from 2005) places Spain at 13th in its ranking, behind only five EU countries, and the trend has been upwards since 1975.

There has been a well-reported drop in inequality between Spain's autonomous regions, in terms of comparisons between the most extreme cases. In 1955, the per-capita income in the richest province (Vizcaya) increased by 4.24, compared with Orense, the poorest province. In 1979, the province of Madrid was 2.4 times richer than Badajoz (Murillo, 1983, 32). Using data that are not completely equivalent in size or geography, according to the LCS the average income per person in Navarra in 2006 was only 1.78 times that of Extremadura.

Poverty indicators include a statistic covering the use and consumption of what are considered basic products. Among the costs specifically mentioned by the LCS, one such indicator would have a reasonably close relation with severe poverty, namely being unable to afford to eat meat or fish at least once every two days. According to this indicator, 2.3% of Spaniards fitted this description in 2007, a figure that was practically the same in 2004 (2.5%). An inability to maintain a suitable temperature affected more people (7.9%, compared with 9.4% in 2004), though this index requires a somewhat more complex explanation (climatic variables, a drop in price of certain domestic appliances, people living alone in large homes, etc.).

Map 1 - *Spain, 2006: Homes in each autonomous region that cannot afford certain costs.*



A: Eating meat or fish at least once every two days.
B: Keeping the home at a suitable temperature.
Source: INE, *Living Conditions Survey, 2007*.

Map 1 shows sharp geographical differences in access to these goods: there is a large area of northern regions (Navarre, Cantabria, Aragón, Asturias and the Basque Country) in a globally enviable position for the southern areas (The Canary Isles, Ceuta and Melilla, Andalusia and Extremadura). Historical differences regarding both income and social inequalities seem to be largely still in place.

According to various studies, the severe poverty threshold historically varies between 25% of the mean income and 40% of the median income. The recent situation in Spain is at best worrying, because the numbers of those affected are not dropping, at least not in recent years. Thus, according to the LCS, the number of people with an income below 40% of the median has grown from 7.5% in 2003 to 7.6% in 2006. Other studies (FOESSA, 2008, 108) report that between 2004 and 2006 the percentage of people below greater poverty thresholds rose slightly (from 3.3% to 3.9% for those below the 30% threshold, and from 2.4% to 2.6% for those who fail to reach even the 25% threshold). In other words, not even the systems in place to guarantee minimum incomes (particularly those aimed at the elderly and the disabled) have been able to prevent the most serious situations, affecting not only immigrants and gypsies, although they are the worst-hit groups.

In the past decade, the rate at which extreme inequalities have been reduced has slowed. Along with Ireland, Spain is an example of a Western European country which has experienced a sharp rise in wealth, but with no drop in the severest levels of poverty. These data should be put in context. Firstly, because the

FOESSA reports indicate a sharp drop in severe poverty in recent decades: from around four million affected people in 1984 to around one and a half million in 1993; in 1979 (Murillo, 1983, 26), 5.7% of the poorest homes generated just 0.51% of the country's income; in 1990-91 (FOESSA, 1994, 322), in the provinces of Salamanca, Ávila and Badajoz, severe poverty affected a quarter of the total population. Furthermore, poverty thresholds are relative indicators. In the decade of growth leading up to 2008, Spanish society went through a period of prosperity that produced a notable increase in the population's spending power. Thus, between 2004 and 2006, real purchasing power (adjusted for inflation) for the average Spaniard (from which any poverty threshold is measured) increased by 4.4%. Many of those currently in extreme poverty are immigrants without appropriate documentation, particularly those who have arrived most recently; this is, therefore, a new, *imported* collective to be added to the old *vicious circles* of traditional poverty, but which do not tend to remain at extreme levels, particularly during times of economic growth. In any case, it is clear that the economic bonanza of the happy years at the beginning of the century has not solved the most extreme inequalities.

Severe exclusion is strongly linked to the gravest situations of poverty, albeit with certain nuances. For some groups, exclusion may be strongly related to poverty, but is not necessarily caused by it. This is the case of minors in foster care and of prison inmates. There are much data on these two collectives in Spain, though none of it is sufficiently useful. Inmate numbers have grown in recent years: according to the Interior Ministry's *Statistical Yearbook*, numbers rose from 41,903 in 1996 to 67,100 in 2007, an increase of 60%, a much higher figure than the increase in population. It may be partially explained by increased immigration, but the native population has also affected the numbers, despite the context of strong wealth and employment creation. The little geographical data available at an autonomous regional level are strongly biased by prison policy and the distribution of prisons, making it impossible to relate the spatial distribution of inmates and the crimes that led to them being imprisoned.

The data on foster care are even more skewed for various aspects, from unaccompanied foreign minors (largely Africans) entering Spain illegally, to how they then spread throughout the country, but they also report a rise in foster care. More directly related to extreme poverty (though also to mental illnesses and immigration, for example), and according to the most recent data available (the 2005 INE *Survey on Homeless People*), there are 21,900 *homeless* people, almost half of whom are foreign (mainly from Africa). Other similar groups include a significant number of AIDS sufferers, the numbers of which tend to go down depending on prevention policies at all levels, and street prostitution, which has a

tendency to increase and which usually involves certain foreign collectives, particularly Nigerian women.

The latest FOESSA report (2009) states that severe exclusion affects many more people in Spain than severe poverty does: 5.3% of the total population. Almost two thirds of them have no pension or alternative benefits; those most affected are the gypsy minorities and some collectives of non-European immigrants. However, there has already been talk of *moneyed* exclusion, particularly among people who live alone, in homes with a disabled person, or immigrants in precarious employment, clearly vulnerable to any potential setbacks.

The same report states that severe exclusion grows much more quickly in medium-sized towns. Without seeking to deny this, we are inclined to think that this may be no more than an issue of how the urban landscape is defined. Much of the poverty in medium-sized cities must undoubtedly be located in towns of rapid growth on the poor outskirts of certain large cities. The process of growing social inequality in large cities has led to the most excluded neighbourhoods in these *false medium-sized cities*.

4. Maintaining strong social inequalities

In recent decades, incomes in Spain have improved to the extent that the country has joined the privileged club of so-called first-world countries, yet remains at a disadvantage in terms of social inequality when compared with fellow first-world nations.

Comparing these data with the LCS for each year, in 2006 the Spanish poverty threshold was at €7,203.30 for single-member households, and €15,126.90 for two adults with two children; this amount would have risen by 14.7% in the past three years. However, the number of households below the poverty line is practically the same, for both men and women. Among EU countries, Spain's situation is not particularly enviable, at least when compared with the countries of the former EU of 15 member states. In Table 1, which shows the recent indices (2007) considered most significant for the Spanish situation with regard to the EU as a whole and to member states with extreme values, the risk of poverty in Spain is one of the highest, not far from the most extreme case (Lithuania), and similar to that of Italy. The situation in Spain is markedly unfavourable among people older than 65 years, particularly if they live alone, and the situation is also worse for workers.

However, the risk of poverty in Spain is similar to that of the whole of the EU, both for people with a low level of education and for single-adult households with children. This situation may be explained by social strengths that have more to do with a past that still influences the present than the latest changes in how society

Table 1 – *European Union, 2007. Rate of risk of poverty and other indicators of living conditions (in percentages)*

Indicator		EU-27	EU-15	Spain	Extreme values	
					Max/country	Min/country
Rate of risk of poverty	TOTAL	16p	17p	20	21/LT	10/CZ/ NL
	<16 years	19(3)	19p	23	25/IT	9/DK
	>65 years	19p	21p	28	51/CY	5/CZ
	>65 years alone	27p	29p	49	75/LV	9/PL
	Workers	8p	8p	11	14/EL	3/CZ
	People with low level of education (1)	14p	14p	15	25/LV	5/NL
	Single-parent families with children	34p	34p	34	54/MT	22/FI
Long-term unemployed		2.6	2.6	2.0	6.6/SK	0.5/CY/ DK
Minors in jobless homes		9.4e	9.2e	5.3	16.7/UK	2.2/SL
Wage gap between men and women		17.4p	18.1p	17.6	30.3/EEp	4.1/ITp
Inequality in income distribution (2)		4.8p	4.9p	5.3	6.5/PTp	3.3/SL

(1) No higher than pre-school, primary or initial secondary. Includes illiterate.
(2) Relationship between income earned by the 20% of the population with highest income (highest quintile) and that earned by the 20% of the population with the lowest income (lowest quintile).
(3) EU-25.
(p) Provisional data
(e) Estimated value

Abbreviations: Cyprus (CY), Denmark (DK), Slovakia (SK), Slovenia (SL), Spain (ES), Estonia (EE), Finland (FI), Greece (EL), Italy (IT), Latvia (LV), Lithuania (LT), Malta (MT), Netherlands (NL), Poland (PL), Portugal (PT), United Kingdom (UK), Czech Republic (CZ), European Union (EU).

Source: EUROSTAT, *Statistics on Income and Living Conditions* (EU-SILC)

is structured. Thus, low levels of education (a remnant of the clearly deficient education system of years gone by) are most prevalent among mature adult workers who lived through the period of stable, lasting work with certain trade-union protections, a situation which deteriorated for subsequent generations. The comparatively low percentages of poverty among single-parent families could be linked to the fact that family networks continue to be solid.

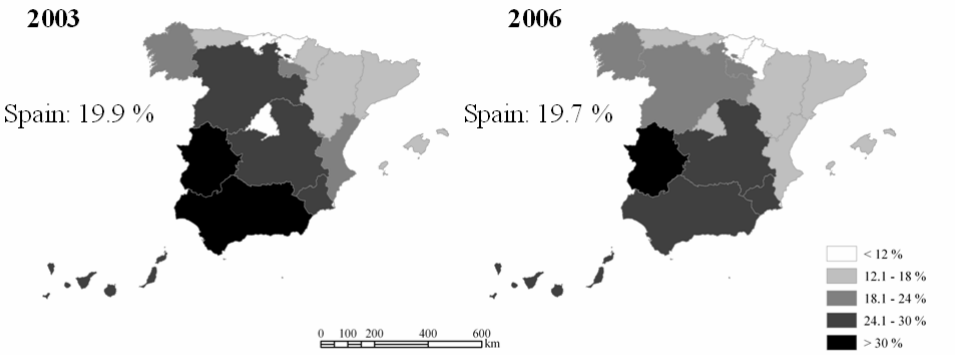
The wage gap between men and women is around the European average, in a country where women's rise in social status is highly visible in aspects such as access to university, involvement in politics and integration in the job market (the current crisis seems to be having a greater effect on men, who form a larger part of the workforce in the most affected sectors, such as industry and construction). The data linked to the job market are favourable: the percentage of long-term unemployed at risk of poverty in 2007 was lower than the European figure and that of the large EU powers (except for the UK); the percentage for underage children in jobless homes was almost half (it was highest in the UK).

Inequality in income distribution is slightly higher in Spain than in the EU as a whole, though slightly lower than the figures for Italy, the UK and Spain's peninsular neighbour Portugal. This inequality is based on significantly lower levels of disposable income than the EU-15 average and on a slightly higher permanent risk of poverty (measured as those currently in poverty who were in the same situation for at least two of the three previous years). Spain is therefore in a subgroup of Mediterranean countries that is defined by a greater degree of inequality and less social welfare than in North-Western Europe, but in better circumstances than the Eastern countries that have recently joined the Union.

Some data may be significant for characterising poverty in Spain in geographic terms, though always at a regional level, which is hardly practical for establishing intervention policies. Maps 2, 3 and 4 show the different figures for those affected by the risk of poverty, homes with incomes of less than €9,000 in 2007, and those who receive basic pensions for old age and disability. Analysed as a whole, certain basic features of the geographic differences in Spanish inequality can be deduced:

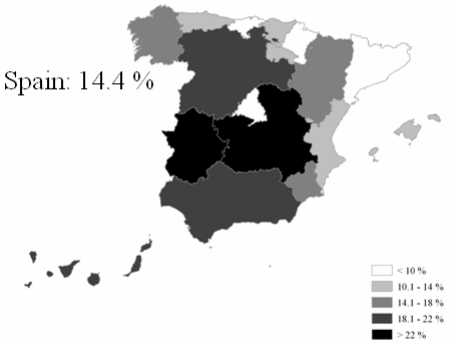
a. There is a series of areas with low scores for these modest income indicators: Navarre, Madrid, the Basque Country and Catalonia, where percentages for the number of people below the poverty threshold and for low-income households are clearly lower than those for Spain as a whole; in these areas, the weighting of those who receive basic pensions is also significantly lower. Navarre may be the most favourable example: according to the 2007 LCS, only 6.3% of its inhabitants live below the relative poverty threshold, and only 7% of its households have an income of less than €9,000; only 5% of the total population receive basic pensions; furthermore, the percentage drop has been the most marked for any region in the past three years; this would mean that Navarre is a comparatively *inclusive* region.

Map 2 - *Spain, 2003-2006: People living below the poverty threshold (*) in each autonomous region.*



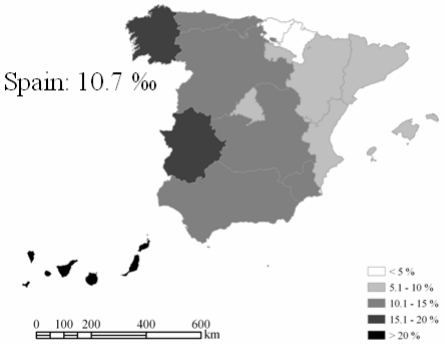
(*) 60% of the median annual income per unit of consumption.
Source: INE, *Living Conditions Survey, 2004* and *2007*.

Map 3 – *Spain, 2006: Households in each region with an income of less than €9,000 per year.*



Source: INE, *Living conditions survey*, 2007.

Map 4 – *Spain, 2006: People receiving old-age and disability pensions (per thousand residents in each autonomous region).*



Source: INE, *Pensions and benefits. Minimum basis benefits, 1994-2006 series.*

b. At the other extreme, Extremadura, Ceuta and Melilla, Castile-La Mancha, Castile-León, Andalusia and the Canary Isles are the regions in the least enviable situation: high percentages of relative poverty and high numbers of low-income inhabitants. These areas have high percentages of inhabitants receiving state pensions, closely linked to the rapid ageing of the population in many rural areas. Recent developments are more disparate. The Canary Isles have worsened slightly over the past three years, even among homes failing to reach certain incomes, despite the positive effect of inflation. Extremadura, the region with the greatest levels of poverty, has worsened in comparative terms during this period. Ceuta and Melilla are in a moderate position with regard to low-income homes, with relatively high percentages of poverty based on a larger family nucleus and a greater inequality gap.

c. The other regions, which have ratings close to the national average, occupy a position halfway between the previous two situations. In general terms, in most of the northern regions (Asturias, Cantabria, Rioja, Aragón), the Balearic Isles and the Region of Valencia, wealth is divided in relatively more equal terms than in Spain as a whole. On the other hand, regions as distant as Galicia and Murcia are much closer to occupying a worrying position, albeit with differences: despite Murcia being a symbol of the so-called *bricks and mortar boom*, relative poverty is high and on the increase; Galicia, meanwhile, known for its rural populations, traditional farming methods (until recently), rapid population ageing, and for receiving more state pension funds than any other Spanish region.

The previous data should be clarified using other variables, as they deal with income and refer to the national average. However, just as incomes vary, so does the cost of living in each region; for example, as of December 2008, the average price for a square metre of property is €2,712, according to La Sociedad de Tasación, S.A., but that is €4,243 in Barcelona, €1,652 in Jaén and only €851 in Villanueva de la Serena (Badajoz), with varying repercussions in terms of the financial commitment required by those with a mortgage to pay. Without reaching such large differences, the cost of living can also differ widely in each province: for example, as Map 3 shows, an income of €9,000 does not represent the same purchasing power in each case.

5. Sociodemographic groups at greatest risk: geographic distribution

The LCS not only sets poverty thresholds, their geographic layout and access to certain levels of consumption (ranging from holidays to air-conditioning at home), it also establishes (using previously established limits) differences between certain variables that help to characterise poverty and, to a lesser degree, social exclusion.

These include their relationship to gender, age, size and type of home, and qualifications.

The aim of this article is not to describe poverty and social exclusion in terms of every single variable to which they are related, but it is necessary to point out how some of them interrelate, because all the factors are intensified by certain circumstances (gender, age, ethnicity, cultural origin, etc.) and have certain influences on each other (Subirats, 2005, 176). For example, the percentages of males and females below the poverty threshold of 40[2] are relatively similar: 7.6% and 7%, respectively. However, there is a much greater difference for single-person households: 7.8% and 3.3%. Household size also explains access or lack of to certain expenses: large families find it more difficult to go on holiday than people on their own, but are more able to cope with unforeseen costs.

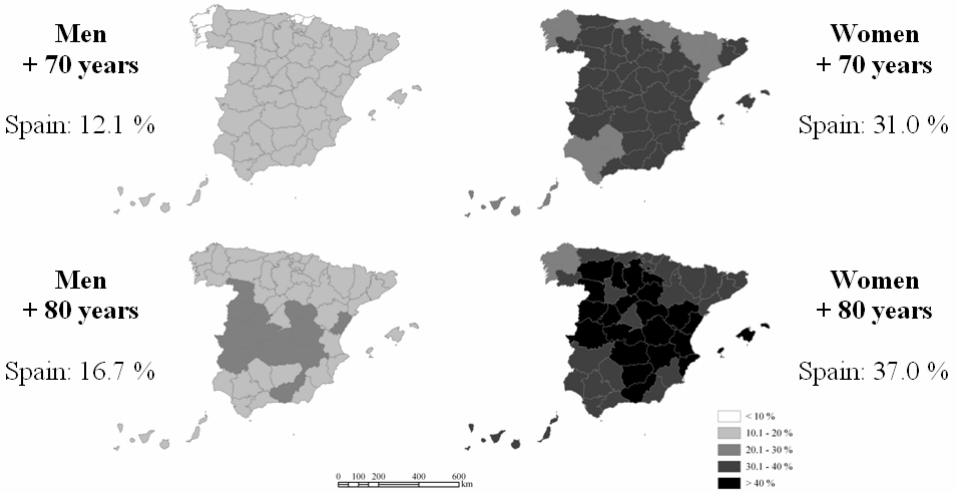
If traditional poverty is linked to misfortune (e.g. the loss of the household's main money earner, illness, losing one's job, etc.), then it is now possible to consider what is known as *new poverty* (Gil, 2002, 34-39), which differs from traditional poverty and is associated with cities, immigrants and the unemployed, who cannot afford a generalised pace of life. Difficulties in accessing housing, single-parent households, mental-health problems and precarious work situations are all linked to relatively high-risk collectives. In general, according to the 2007 LCS, the risk of poverty increases in homes where people live alone, for people aged over 65, single parents with young children and with no more than a primary level of education, the unemployed and the economically inactive. However, the risk exists in almost any collective: 10.6% of those in work are in a situation of poverty, as are 8.1% of people with a higher level of education; these percentages are higher than those of Navarre as a whole (6.3%). It would thus seem significant to point out the geographic component of poverty and, perhaps even more so (though the figures are less reliable), of social exclusion.

5.1 *Living alone and illiteracy among the aged*

Map 5 measures the influence of the relationships between factors that increase the risk of poverty and the strong regional variations within these factors. Gender, age and size of household influence poverty, but so do differences between regions, to a great extent. We have considered that, of all the factors given by the LCS that characterise poverty, the greatest correlation exists with regard to old people who live alone.

[2] Those below 40% of the median national income.

Map 5 – *Spain, 2001: People who live alone, divided into women and men over the age of 70 and 80, by provinces.*



Source, INE: *Spanish Population Census*, 2001.

Combined with this factor (illustrated by considering people over 70 and 80 years old on separate maps), we have also differentiated by gender and for provinces, to determine the increased risk. The differences are clear: for women over 70 the figure was 31% in 2001, compared with 12% for men; among women aged over 80, who have greater problems of personal autonomy and need more help, the figure rose to 37%, with figures as high as 48% in provinces such as Ávila and Cuenca, which are known to have suffered from a massive rural exodus in previous decades, and are currently experiencing strong population ageing. One fact may fully explain this strong inequality between sexes when considering living alone in old age: in Pontevedra, where the fewest number of women over 80 live alone, the situation affects 25.5% of the total, which is higher than the figure for males in any province. These differences can be explained by many demographic and sociological factors: the fact that the vast majority of women live longer than their partners (they are generally younger and their life expectancy is higher), their greater traditional link to work around the house helps them to survive more easily without resorting to family help, etc.

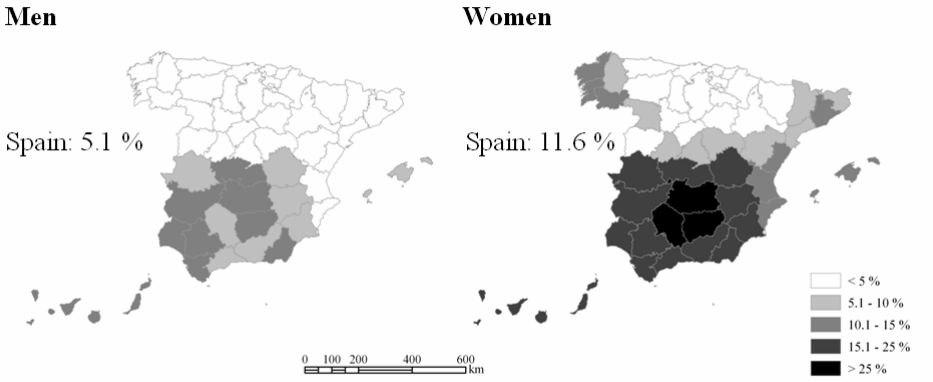
The relationship between people living alone and poverty is strengthened by the fact that the family structure is still the most solid care network for citizens and an effective buffer against serious difficulties. In the elderly, leaving the working world is usually accompanied by a drop in income. The blanket of benefits available for the elderly and the increase in minimum pensions has lessened their

risk of poverty, but some provisions are still clearly insufficient, such as the basic pension for those with no right to any other, which in 2008 was only €5,156.62 per year. It is understandable that women are predominantly among those who find themselves in this situation, having lived through a period when regular employment for women was rare.

Another risk factor inherited from the past that is currently affecting elderly generations is illiteracy. According to the 2001 *Population Census*, 2.4% of Spain's population aged 10 years old or over was affected by illiteracy. However, the correlation with old age (in a country where in the past the right to basic education was more theoretical than real) is very high: in only two generations, Spain has gone from being a country with an excessive number of illiterate women to one in which there are more young females than young males with top qualifications. Today, illiteracy among the elderly is no longer a risk for their job possibilities, but given that there is a clear degree of poverty being passed on from one generation to the next, illiteracy among the elderly can provide warnings of where there may be gaps in education and other related difficulties. The geographic distribution of illiteracy among the elderly as shown in Map 6 explains many of the strong inequalities that were characteristic of Spain in the past and which to a certain extent still affect society to this day.

According to the map, the percentages of illiterate people over the age of 70 are much higher than for the population as a whole, with a strong negative bias for women (11.6% compared with 5.1% for men). The reason is clear: the serious shortcomings in Spanish education prior to 1960, and the difference in how the two genders were taught, which clearly discriminated against girls. However, the regional differences are even greater, with provinces such as Segovia, where only 1.9% of women and 0.9% of men of this age are illiterate, and others such as Jaén, where the figures exceed 29% and 13%, respectively. In general, three clearly defined regions can be established with regard to illiteracy among the elderly:

Map 6– *Spain, 2001: Illiterates by gender, among people aged over 80 in each province.*



Source, INE: *Spanish Population Census*, 2001.

a. The traditional area of illiteracy in Spain, which affects more than 18% of the elderly: all the provinces in Andalusia, Extremadura, the Canary Isles and Murcia, the cities of Ceuta and Melilla, and the provinces of Ciudad Real and Toledo in Castile-La Mancha. Perhaps no other classification divides the north and south of Spain so clearly, reflecting the traditions of former farming methods and ways of life, involving large estates and an atavistic inequality linked to land dominance. In all of these (with the slight exception of Málaga), more than 10% of the males over 70 years old are illiterate.

b. On the opposite side is a large area of the middle and north of Spain, made up of Asturias, Cantabria, the Basque Country, Navarre and La Rioja, as well as the northern provinces of Aragón (Huesca and Zaragoza) and almost all of Castile-León (except Ávila and Zamora), where fewer than 5% of women, and fewer than 2% of men, are illiterate. This is an area characterised by traditional smallholdings and where schooling for both sexes was a reality during the early decades of the 20th century.

c. These are provinces with relatively low levels of illiteracy, including Madrid and Catalonia, as well as some of the provinces neighbouring those mentioned above (Ávila, Guadalajara, Zamora, Teruel and Lugo), where fewer than 10% of elderly women are illiterate. Levels increased in Madrid and Cataluña with the arrival of those from rural areas in the mid-20th century. The rest of the provinces

share many of the characteristics of group b, although in general these are poorer areas than those in group b.

d. This fourth group has high levels of illiteracy, though not as pronounced as those in the southern part of Spain. These include the Region of Valencia and the Balearic Isles (with strong inequalities in their traditional rural structure); also included is most of Galicia (except for Lugo), which is a traditionally poor region made of scattered smallholdings, and the provinces of Albacete and Cuenca in Castile-La Mancha, which are transitional areas between the large farmed estates and regions with mainly smallholdings.

With all these clarifications, the map of illiteracy among the elderly is not dissimilar to that of the current impact of relative and severe poverty: it is an example of how the relationship between poverty and the level of education has helped or hindered the changes and improvements to society and the economy. It also goes some way to explaining the previously mentioned notion of poverty passed on from one generation to the next.

5.2 *The gypsy population*

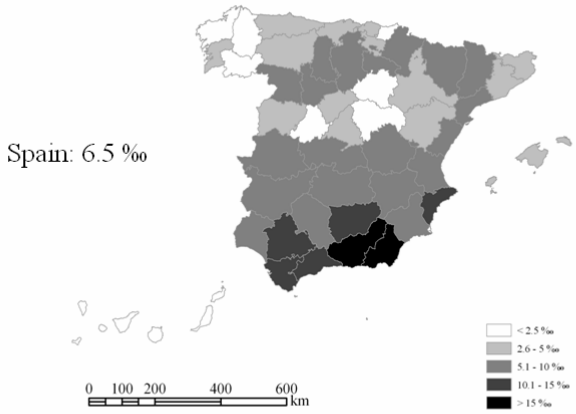
Gypsies have repeatedly been the hardest hit minority in terms of poverty and social exclusion. Today they continue to face the biggest problems in terms of integration, and gypsies are still the collective with the lowest prestige in Spanish society: in late 2007, only 36.9% of the population had much or some sympathy towards them, compared with 51.7% who had little or none (CIS, 2007); the positive percentage was less than that reported for Muslims and much lower than that given for immigrants. These are prejudices which, as Guiddens points out (2007, 472), tend to be based on stereotypes and fixed, unwavering views towards a group, and seem to affect ethnic minorities the most, by discriminating against them and depriving them of the opportunities enjoyed by others.

The LCS makes no direct reference to the gypsy population. The same is true for censuses, municipal registers, civil registries and all other statistical or official ongoing population counts in Spain. Only specific studies and estimates give any assessment of this the most socially excluded collective, with up to 12% of those in a situation of severe exclusion in Spain, in a community that probably makes up no more than 2% of the total population (FOESSA, 2008, 202). Based on regional studies, the FOESSA report puts the gypsy community at around 970,000 people; it also states that, unlike the stereotypes usually applied, the gypsy community has employment rates above the Spanish average, as well as unemployment rates slightly above the average. If the situation seems quite normal, their access to work is marked by under-employment and temporary work, with high levels of self-

employment in family businesses, and only 7.4% of the active gypsy population are estimated to be in a permanent salaried job. If an equivalence were established with regulated employment, real unemployment levels could be as high as 38% (FOESSA, 2003). There appears to have been considerable improvement in education levels in terms of the compensatory education programme implemented by the education authorities at the end of the 1980s: no longer are there figures such as 22.3% of gypsies who went no further than initial primary schooling, or 53.9% who received no more than a secondary education (MEC, 1989); however, poor education levels, falling behind and absenteeism from pre-adolescence all affect a significant number of gypsy schoolchildren.

The material conditions of the gypsy population are set out in the *Map showing the gypsy community and housing in Spain*, which was produced in 2007 by the Gypsy Secretariat. This organisation estimates there to be 92,770 gypsy households, which would not seem to fit with the figures given by the FOESSA report, unless we accept an average of 10.45 persons per household, which is highly unlikely, unless the typical ancestral stereotypes are to be believed.

Map 7 – *Spain, 2007: Gipsy houseshold in each province (in terms of the total number of main family homes*).*



(*) Calculated using the number of housesholds in the 2001 *Population and Households Census.*
Source: Gipsy Secretariat: *Map showing the gipsy community and housing, 2007.*

With the exception of the two conflicting sets of data (which does not mean that both studies are unacceptable), the *Map* states that half of the gypsy community lives in subsidised housing and sets out the extent of certain problems; for example, 24.8% of the buildings are in a poor state of repair, in 19% of cases the streets have

deficiencies, in 16% of cases there are problems with communication links, and 24.7% are in a situation of particular vulnerability. In total, 6,223 homes were in a state of serious deterioration, 500 were the equivalent of shanties or makeshift homes, 196 were in buildings designed for other purposes, and 314 were mobile homes. Around 5% of the total were settlements set apart from the rest of the town. Map 7 shows data supplied by the Gypsy Secretariat on the different relative weighting of the gypsy collective in Spain, with high numbers in the south-eastern provinces of the country, particularly Almería (with more than 3% of the total of main households) and Granada; on the other hand, very few gypsies live in the Canary Isles, Ceuta and Melilla, Galicia and the rural areas of Castile around Madrid to the north.

6. Foreign immigration: the major new challenge for social integration

The major sociodemographic transformation that Spain has undergone in the past decade is related to immigration, with people arriving from various origins – mainly third-world countries, attracted by strong job growth, employment gaps as a result of the segmentation of the job market, and also by the relative ease of entering and remaining in the country; many also arrive from the North-Atlantic countries of Europe to settle in areas with a favourable climate and abundant tourist infrastructure. In total, the number of foreign residents in Spain has gone from just over 600,000 in 1998 to more than five million ten years later: in January 2008 they made up 11.4% of the country's registered population.

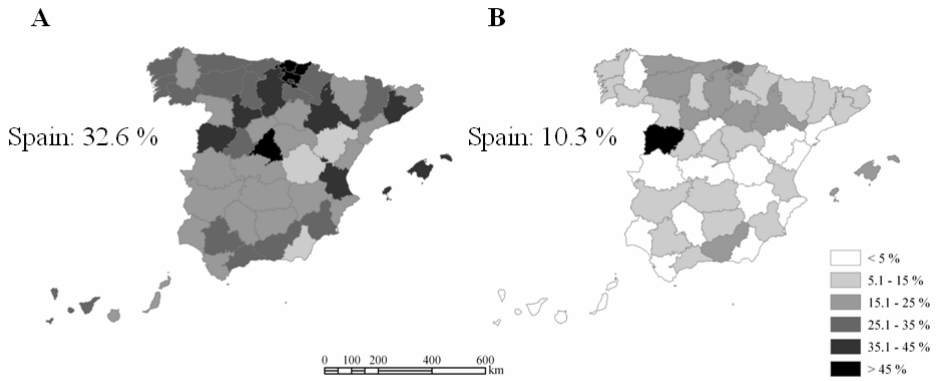
Many foreign immigrants who come to Spain looking for work make up the collectives associated with *new poverty*. Of course, not all immigrants, or even a majority of them, are in a situation of poverty or social exclusion, although it is hard to give numbers because the LCS provides no data on immigrants or foreigners. The latest FOESSA report (2009) states that only a small number of immigrants must be in a situation of severe exclusion – around 14% of this highly heterogeneous collective. It is hoped that the high correlation between immigration, poverty and exclusion may fall over time; it should be remembered that, for many, the process of adaptation to the country involves living as undocumented citizens, and they are often forced to take on jobs abandoned or not accepted by the native population. One way to consider this collective is as a group of people whose knowledge of the language and their new surroundings improves every day; immigrants tend to be decisive people (they are, among other achievements, capable of taking on a project of migration) and almost always have an above-average level of training or education for their country of origin.

This new poverty as represented by foreign immigrant workers includes some worrying features, because high levels of poverty tend to go hand in hand with high rates of employment activity. Predominant among this group are young adults, who have the greatest potential in terms of work possibilities, so their levels of employment are higher than the average for Spain: in many cases, several income-earners live in the same household; they are more willing to work overtime and difficult hours, such as at night and weekends, etc. Despite this, they are the collective with the highest percentage of people in employment living below the poverty threshold. Many work in the worst-paid, most precarious and most instable jobs, with many (particularly women) working in barely legal circumstances (Gozálvez, 2008, 176-179). This is considered by some as a new, *exogamic* kind of inequality (compared with traditional *endogamic* inequality), in which individuals from other cultures have accepted unequal conditions which until recently the native population fought hard to eradicate, and who now would appear to have suspended their previous beliefs (Gil, 2002).

Traditional European society had two instruments of social integration, among others: schools, considered to be a centre of homogenisation and a *social elevator*; and living in shared, ghetto-free neighbourhoods. How do these affect the new residents of Spain? Clearly this depends on each national collective, and even on each place of settlement: things are not the same for the *false tourist* from Norway as they are for the immigrant from Argentina or the Gambia; areas of intensive agriculture have little in common with manufacturing environments. Maps 8 and 9 analyse the question with regard to immigration from African countries (both North and sub-Saharan Africa), which appears to be the group most likely to face rejection[3]. Map 8 shows the difference in schooling of African children, who have the greatest adaptation problems, from linguistic difficulties to rejection by some parents. These are children who face all manner of de facto problems to enrol in private and charter-based schools, even though the latter are funded by the State. With the exception of Salamanca, in the whole of Spain the percentage of African pupils enrolled in private schools is much lower than what would be an equitable figure.

[3] The previously mentioned study by the CIS (Sociological Research Centre) on discrimination and its perception, published in December 2007, considered that Spanish people have the least sympathy for people of the Muslim faith, after gypsies, although the percentage for those who had no sympathy whatsoever for them was slightly higher. Although Muslims and Africans are not exactly the same collective, there is a very high correlation of perception between the two groups in Spain.

Map 8 – *Spain, 2006-07: Non-university students enrolled in private schools.*



A: Percentage of all pupils.
B: Percentage of African pupils.
Source: MEC, *Statistics of non-university teaching. Detailed results 2006-07.*

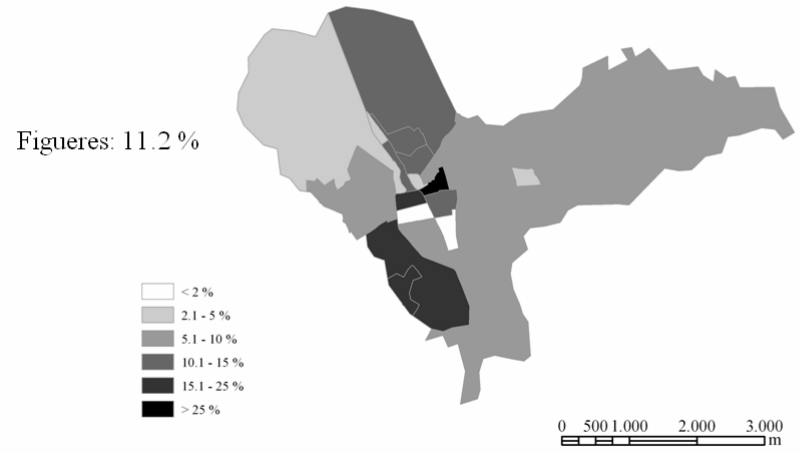
The situation is more extreme where private schools are less prevalent, i.e. where they are more elitist in nature; thus, using data from the 2006-07 school year, in Melilla, private schools counted for 16.1% of the city's schoolchildren, but only 0.8% of foreign African children; in the province of Almería, these figures were 13.6% and 0.8%; and in Tenerife, 25.6% and 2.3%.

Where steps have been taken to correct excessive concentrations, the result is far from balanced: in Barcelona, 43% of children go to private schools, but only 14% of Africans; in Madrid, these figures are 46% and 12%. In practice, the public network takes the lion's share of schooling children with greater difficulties. There is a risk of the public education system becoming segregated between state-run schools with high levels of immigration and a very limited blend of children from different social classes, and charter-based schools used by those who do not want their children to mix with the others. As a result, schools would increasingly lose their roles as places of equal opportunities (Valero, 2002,167).

In large cities, immigration has sharpened what Guiddens (2007, 857) calls the *geography of centrality and marginality*, in which opulence and poverty coexist with hardly any real contact. In many places, immigrants tend to gather in certain areas, such as historic town centres (Valero, 2008, 59-61), which were unable to adapt to modern times and modern needs, or certain neighbourhoods that grew up as a result of the rural exodus of the 1960s, where the housing that was built was of poor quality. This is true not only of the largest cities. We can see a specific

distribution in the case of Figueres, a small town in Catalonia (Alt Empordà, in the province of Girona), where 11.2% of the population are immigrants, a figure close to the national average. Map 9 shows strong differences in terms of where immigrants have settled in the town's various census-registered sections: in section 3.4, 46% of residents are African (to many of whom would be added their children, born in Spain and some with Spanish nationality); in sections 1.1 and 4.7, fewer than 2% of the residents are African. The two areas are only a few minutes' walk apart.

Map 9 – *Figueres (Girona), 2008: Percentage of Africans compared with the total population of each of the town's census sections.*



Source: INE, *Municipal Register of Inhabitants*, 1-1-2008

This coexistence in very close physical but socially distant areas helps to foster negative stereotypes, such as the association of immigration with delinquency, a view which is widespread among large sectors of the Spanish population. This statement is, like so many, only a half truth: according to the Ministry of the Interior, in late 2007, foreigners made up 34% of the Spanish prison population, a figure that is interestingly, very close to that of severe social exclusion among foreigners. Their number has tripled since 1997, but the number of foreigners in Spain has multiplied by more than eight. In other words, as the number of documented immigrants has increased and it has been easier to find legal jobs, the percentage of immigrant inmates has dropped. However, times of recession could see an increase in the fact that large parts of the Spanish population reject the competitiveness in the job market that foreigners represent, and consider

immigration to be a problem, both of which were true even during the boom years (Gozálvez, 2006, 45-47), as the number of criminal acts increases linked to the increase in precarious situations.

Finally, another question in which it is easy to determine how poverty and exclusion feed into each other is in terms of family reunification. Those lacking the possibility to gain access to free housing, a stable job and a minimum income (i.e. those who are affected in some way by poverty) also find it impossible to successfully secure any kind of family reunification, and they are therefore at greater risk of falling into a situation of permanent exclusion. By analysing the demographic pyramids of some African collectives that came to Spain early on, such as the sub-Saharans who settled in areas between Barcelona and the Pyrenees, there are significant groups of single males with no family, living together in groups as a lost generation and in a clear situation of social exclusion.

7. Conclusions

There has been a drop in extreme poverty in Spain over recent decades, in accordance with the country's transformation in terms of its social, economic and regional characteristics.

However, relative poverty still exceeds the average for Europe and remains an unavoidable challenge, particularly because the rapid economic development of the past decade has brought no clear reduction in social inequalities.

In general, the increase in minimum pensions and other benefits has reduced the severe poverty that some collectives such as the elderly were facing. This has helped to improve the situation in rural areas, which have the highest rates of population ageing and a large element of traditional poverty.

Despite these transformations, in general terms there are still two well-defined areas of the country, in terms of both the percentage of the population in a situation of severe poverty and those living below various poverty thresholds: a southern area based on Extremadura and Andalusia, and a northern region, centred round Navarre.

The interrelationship of factors such as gender, old age and single-family homes establishes one of the demographic collectives at high risk of social exclusion. On the other hand, areas of the country that were traditionally affected by high levels of illiteracy tend, to a certain degree, to also define the geography of poverty.

Two well-defined collectives, though not ones that are directly reported in the LCS (gypsies and foreign immigrants), are the two big examples of where poverty and social exclusion combine in Spain. Among other discriminatory features, these are the groups that suffer most from prejudice and negative stereotypes, as well as

experiencing problems in schooling and habitat segregation. Gypsies are always the clearest example of the most traditional, almost atavistic, kind of poverty and social exclusion. Economic immigrants, meanwhile (whose numbers have multiplied by more than eight in the past decade), are the collective most closely linked to what is known as *New Poverty*.

References

Alonso F.J. 1995. *Reflexiones en torno a la investigaciones sobre pobreza y marginación en España*. In: Álvarez F. y otros: Desigualdad y pobreza hoy, pp.89-105, Talasa, Madrid.
CIS (Centro de Investigaciones Sociológicas). 2007. *Estudio sobre la discriminación y su percepción*, (www.cis.es).
FOESSA, Fundación. 1994. V Informe sociológico sobre la situación social en España, Fundación FOESSA. (www.foessa.org).
FOESSA, Fundación. 2009. *VI Informe sobre exclusión y desarrollo social en España*, (www.foessa.org).
Gil, F. 2002. *La exclusión social,* Ariel, Barcelona.
Giner S., Espinosa E.L. y Torres C. (eds). 2006. *Diccionario de Sociología*, Alianza, Madrid.
Gozálvez V. 2006. *Los indicadores de integración de los inmigrantes en España*. In: Gozálvez V. (dir) La inmigración extranjera como desafío y esperanza, pp. 41-58, Departamento de Geografía Humana.Universidad de Alicante.
Gozálvez V. 2008. *La diversificación de los extranjeros en España: sus implicaciones*, Quaderni, nº 35, pp 151-194, Dipartimento per lo Studio delle società mediterranee. Università degli Studi di Bari.
Guiddens A. 2007 *Manual de sociología,* Alianza, Madrid.
PNUD (Programa de Naciones Unidas para el Desarrollo). 2009. *Informe sobre Desarrollo Humano, 2007-08*, (www.undp.org/spanish/).
Ministerio de Educación, Política Social y Deporte. 2008. *Estadísticas de enseñanzas no universitarias. Resultados detallados del curso 2006-07.*
Ministerio de Educación y Ciencia. 1989. *Memoria de educación compensatoria del curso 1988-1989.*
Ministerio del Interior. 2008. *Anuario Estadístico* (www.mir.es).
Murillo, F. y otros. 1983. *Informe sociológico sobre el cambio social en España, 1975-1983*, Fundación FOESSA, Madrid.
Secretariado Gitano, Fundación. 2008. *Mapa sobre vivienda y comunidad gitana en España, 2007.* (www.gitanos.org)
Subirats J. (dir). 2005. *Análisis de los factores de exclusión social*. Documentos de trabajo, nº 4, Fundación BBVA.
Valero, J.R. 2002. *Hacia una escuela multicultural: la presencia creciente de los inmigrantes en el sistema educativo español*, Cuadernos de Geografía, nº 72, pp. 152-172, Universidad de Valencia.

Valero J.R. 2008. *La inmigración en los centros históricos de las ciudades*. In: Valero J. (ed.) La inmigración en los centros históricos, pp. 45-73, Departamento de Geografía Humana. Universidad de Alicante.
Vázquez J.M. (dir). 1990. *Estudio sociológico: los gitanos españoles en 1978*, Secretariado Gitano, Madrid.

SUMMARY

The notable economic development that Spain has experienced in recent years has not produced a parallel drop in social inequalities. The increase in minimum pensions and other social benefits has reduced severe poverty and improved life for elderly people and in rural areas, but strong inequalities still exist. In the southern regions of mainland Spain, the risk of poverty increases for single-person households and elderly women. Gypsies are the collective that traditionally is most closely linked to social exclusion, and foreign immigrants are the largest group among the *new* poor.

Vicente GOZÁLVEZ PÉREZ, Human Geography Department of the University of Alicante (Spain).
José Ramón VALERO ESCANDELL, Human Geography Department of the University of Alicante (Spain).

Rivista Italiana di Economia Demografia e Statistica Volume LXIII nn. 1-2 – Gennaio-Giugno 2009

GROWING COHESIVE SOCIETIES: EXPLORING THE LINK BETWEEN SOCIAL EXCLUSION AND ACTIVE CITIZENSHIP

Massimiliano Mascherini, Andrea Saltelli

1. Introduction

Facilitating Active citizenship is one of the European Commission's strategies for increasing social cohesion and reducing the democratic deficit across Europe within the context of the wider Lisbon process. In this regard indicators have been requested by member states (Council 2005 and Council 2007) then developed by CRELL (Hoskins et al 2006, Hoskins et al 2008 and Hoskins and Mascherini 2009) and used within the European Commission Progress reports on the Lisbon process (European Commission 2007 and European Commission 2008). The next research step, towards deepening the understanding of this phenomenon and towards providing an evidence base for policy development, was to identify the socio-demographic characteristics and determinants of active citizens and those who for one reason or another participate much less. This report provides a detailed identikit of the active citizen from 2002 across 14 European countries Austrian, Belgium, Germany, Denmark, Spain, Finland, United Kingdom, Greece, Italy, Luxemburg, Netherlands, Norway, Portugal, Sweden (the complete dataset available for this research is only available for the majority of old member states of the European Union and European Economic Area).

In this context, the aim of the report is to deepen the understanding of Active Citizenship by identifying the determinants of Active Citizenship through the application of a multilevel model that examines both the individual level and national level characteristics. Hoskins and Mascherini (2009) presented a composite indicator to measure Active Citizenship based on 61 basic indicators drawn from the 2002 European Social Survey data. Following this framework, individual level analysis is carried out using socio-demographic and behavioral variables of gender, occupation, income, age, religion and use of media of active citizens. On a national level it provides an analysis of the contextual features of the country which enhance active citizenship such as; GDP, income equality, national averages of education and religious diversity. This research also enables a greater understanding of who is much less active.

Research in the field of political participation has shown that in the US (Verba,

Schlozman and Brady, 1995) and across 62 diverse countries in the world (Norris 2002) that the individual characteristics of gender, ethnicity and social class have not been found to be significant predictors of political participation after controlling for education, occupation and social and economic status. Norris (2002) across the 62 diverse countries and Lauglo and Oia (2002) in Norway found that age was a significant factor with participation increasing with age and in the case of Norris's research, she found that the middle aged participated the most. Verba, Slozman and Brady (1995), found that family income is a predictor of political voice and influence. Education across the years has been identified as the single most important predictor of different forms of political participation (Dee 2004, Finkel 2003, Print 2007, Galston 2001, Verba, Schlozsm and Hoskins et al 2008).

The effect of the media and news has had conflicting results as Semetko 2007 noted in a review of this literature for voter turn out. She highlighted that there was equal evidence of media increasing cynicism and reducing engagement as there was for it increasing the levels of citizen's involvement, trust and efficacy. Based on the previous literature, what we can expect to see is that age, education and wealth are the key features of active citizenship. In terms of age we would expect to see the middle age participate more. Concerning education and wealth the more you have the more we would expect that people participate.

The potential barriers to active citizenship have been described by Hoskins et al (2008) as 'financial concerns (e.g. paying subscriptions to be a party member), in terms of spare time (e.g. if an individual is both working and looking after a family), geographical location (e.g. in the countryside without good public transport) and information (e.g. being part of networks that keep you informed).' Verba, Slozman and Brady 1995 categorized the barriers that they had found from their research into 3 major reasons for not being able to participate, 1) they can't, due to a lack of money, time and skills, 2) they don't want to, due to no interest, they think it makes no difference and a limited knowledge of process 3) nobody asked (they lacked information). They suggest that the extent that these factors influence the levels of participation depends on which forms of participation are under discussion. This approach that is used predominantly on research on elections, does not help to explain why so many people actually vote. From this research we would expect to see that wealth, amount of free time, geographical location, information from various media sources and involvement in social networks would be crucial to whether people are active citizens.

In this paper, we identify which socio-demographic features are critical to active citizenship in 14 European countries and which social groups are more isolated and participate much less.

This paper is organized into three sections. Section 2 describes the active citizenship composite indicator and in Section 3 possible socio-economic and

behavioural determinants of Active Citizenship through individual data and multilevel analysis are deeply investigated. The results are finally described, commented upon and conclusions drawn. Finally issues to be addressed by further research are presented.

2. The Active Citizenship Composite Indicator

Building on the foundations of Marshall (1950) in terms of rights and obligations of citizenship and Verba and Nie (1972) in terms of participatory and influential action, Hoskins and Mascherini (2009) defined active citizenship as:

> *"Participation in civil society, community and/or political life, characterised by mutual respect and non-violence and in accordance with human rights and democracy."*
> (Hoskins, 2006)

As can be seen within this definition, Active citizenship incorporates a wide spread of participatory activities containing political action, participatory democracy and civil society and community support. However, and in our view correctly, action alone is not considered active citizenship, the examples of Nazi Germany or Communist Europe can show mass participation without necessarily democratic or beneficial consequences. Instead participation is incorporated with democratic values, mutual respect and human rights. Thus what we are attempting to measure is value based participation. The difference between this concept and social capital is that the emphasis is placed on the societal outcomes of democracy and social cohesion and not on the benefits to the individual from participation. For further details on the conceptual development of active citizenship we address the reader to Hoskins and Mascherini, 2009.

After defining the concept, Hoskins and Mascherini, 2009 based the operational model of active citizenship on four measurable and distinct dimensions of Protest and social change, Community life, Representative democracy and Democratic values. The dimension on Protest and Social change is comprised of four components. The first component is protest activities which is a combination of 5 indicators: signing a petition, taking part in a lawful demonstration, boycotting products and contacting a politician. The next 3 components are three types of organizations; human rights organisations, trade unions and environmental organisations. Each of these components is comprised of four indicators on membership, participation activities, donating money and voluntary work. The Community life dimension is comprised of seven components. Six of these are

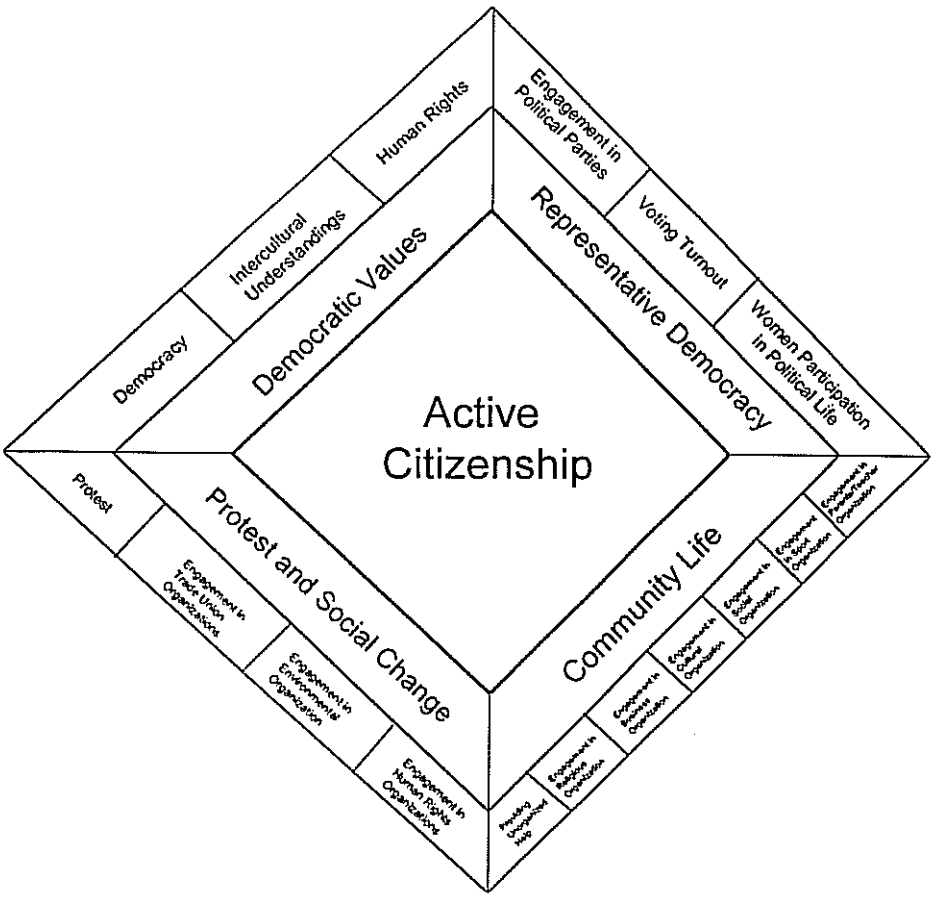
community organisations: religious, business, cultural, social, sport and parent-teacher organisations. These 6 components contain 4 indicators each on membership, participation activities, donating money and voluntary work. The 7th component is a single indicator on unorganized help. The dimension Representative democracy is built from 3 sub-dimensions; engagement in political parties, voter turnout and participation of women in political life. The subdimension on engagement in political parties contains 4 indicators on membership, participation, donating money or voluntary work for political parties. The subdimension on voter turn out contains two indicators on voting, one on the national elections and one on European elections. The third sub-dimension is comprised of one indicator on the percentage of women in national parliaments. The fourth dimension is called Democratic values and consists of 3 sub-domains: democracy, intercultural understanding and human rights. The democracy sub-domain is comprised of 5 indicators on Democratic Values asked in relationship to citizenship activities. The intercultural sub-dimension contains 3 indicators on immigration. The human rights sub-dimension is comprised of 3 indicators on human rights in relationship to law and rights of migrants.

The operational model adopted to measure Active Citizenship is described in figure below. For the complete list of indicators we address the reader to and Hoskins and Mascherini 2009.

2.1 *Data and Methods*

In the field of active citizenship availability of data is a serious problem. Not all dimensions are sufficiently covered and multi-annual data are generally not available. For example, there are limited data available on more informal and less conventional methods of participation, which have been seen to rise in recent years and which are often more culturally specific. Where possible non-conventional participation such as ethical consumption and unorganized participation have been included in the model, but the data for traditional forms of participation are more plentiful and easier to access from survey data.

Figure 1 – *The Structure of the Actvie Citizenship Composite Indicator.*



With this in mind, the selection of indicators for the composite measure of active citizenship has been based mostly upon one source of data, which helps to maximize the comparability of the indicators. The source of data chosen was the European Social Survey (http://www.europeansocialsurvey.org/) which ran a specific module on citizenship in 2002. The European Social Survey (ESS) aimed to be representative of all residents among the population aged 15 years and above in each participating country. The size and the quality of the sample make the country coverage of Europe in the ESS data reasonably good, with 19 European countries, including 18 EU member states, providing sufficient quality of data.

Overall, the Active Citizenship Composite Indicator is based on a list of 61 basic indicators. As stated above, most of these indicators use individual data collected in the European Social Survey of 2002. In addition, voter turnout at national and European elections has also been considered, as well as the proportion of women in national parliaments. In order to complete the dataset, one missing value has been imputed for Norway.. The list of the 19 countries included in the analysis is given in table 1 below. The list of the basic indicators can be found in Hoskins and Mascherini 2009.

Table 1 – *List of countries included in the Active Citizenship Composite Indicator.*

List of Countries			
Austria	*Netherlands*	*Finland*	*Slovenia*
Italy	*Denmark*	*Portugal*	*Greece*
Belgium	*Norway*	*France*	*Ireland*
Luxembourg	*Spain*	*Sweden*	*Hungary*
Germany	*Poland*	*United Kingdom*	

Nardo et al. (2005) define a composite indicator as "a mathematical combination of individual indicators that represent different dimensions of a concept whose description is the objective of the analysis". Following this logic, here we summarize the concept of active citizenship into one number, a composite indicator, which encompasses different dimensions.

We built the composite indicators following the methodological guidelines given by Nardo et al. (2005). In this paper the different phases of the construction process of the composite indicators are just sketched and we address the reader to Hoskins and Mascherini, 2008 for details and wider description.

Given the structure of the Active Citizenship Composite Indicator shown in figure 1, the composite indicator is a weighted sum of the indices computed for the four dimensions D_i (Representative Democracy, Protest and social change, Community, Democratic Values) with weights w_i. The indices of each dimension D_i is then a linear weighted sum of of the sub-dimension indices SD_{ij}. with weights

w_j^*. Finally, each sub-dimension index SD_{ij} is a linear weighted aggregation of the s_{ij} normalised sub-indicators $I_{h_{i,j}c}$ with weights $w_{h_{i,j}}^{\#}$ The integration of the different equations into one gives the general formula for the Active Citizenship Composite Indicator:

$$Y_c = \sum_{i=1}^{4} w_i \sum_{j=1}^{k_i} w_j^* \sum_{h_{ij}=1}^{s_{ij}} w_h^{\#} I_{h_{i,j}c}$$

Having defined the aggregation rule of the composite indicator, the construction and evaluation of the composite indicator (CI) involve several steps. In the next step the variables must be standardized and the weighting scheme for the indicators specified. Due to the fact that the 61 basic indicators have been constructed using different scales, a standardization process is needed before the data for the different indicators can be aggregated. Different standardization techniques are available for this (Nardo et al., 2005). The basic standardization technique that has been applied is the well known z score approach in which for each basic indicator , $x_{m,n}$, the average across countries and the standard deviation across countries are calculated. The normalization formula is:

$$I_{m,n} = \frac{x_{m,n} - \bar{x}_m}{\sigma_{x_m}}.$$

After the standardization process, the data have then been transformed to ensure that for each indicator a higher score would point to a better performance. This step was clearly necessary to make a meaningful aggregation of the different indicators. Based on the Active Citizenship Composite Indicator structure the weights were assigned after the consultation of experts in the field of active citizenship. This was done in order to assign different weights to the various dimensions on the basis of experts judgment which was elicited with a survey designed following the budget allocation approach. In order to permit the elicitation of the experts' judgment, on February 2007 we distributed a questionnaire to 27 leading experts on Active Citizenship. All of the people contacted for participating in the survey had been established as researchers or key experts in the field of the Active Citizenship domain and for this reason they were considered experts. In particular, the participants to the survey belong to 4 different areas of expertise: sociologists, political scientists, policy makers and educationalists.

The questionnaire was designed following the budget allocation approach, that is a participatory method in which experts are given a "budget" of N points (in our

case 100), to be distributed over a number of sub-indicators, paying more for those indicators whose importance they want to stress. (Moldan and Billharz, 1997). For each expert, the weights of the basic indicators were computed by a linear combination of normalized values of the median of the distribution of the weights assigned to dimensions and sub dimensions. For a detailed description of the computation of the weights and the experts' elicitation process we address the reader to Mascherini and Hoskins, 2008. Finally a consistent sensitivity analysis was performed in order to show the robustness of the composite indicator which is not affected by the assumption made in the construction process.

Moreover in Hoskins et al. 2006 and Hoskins and Mascherini, 2008 a consistent sensitivity analysis was performed in order to successfully show the robustness of the composite indicator that is not affected by the assumption made in the construction process.

The composite indicator is then computed on the basis of the weights elicited by the experts. For each expert, the composite indicator is computed once for all countries. The score assigned to each country corresponds to the median of the distribution of the scores assigned to that country by all the experts.

Overall, it can be seen that the Nordic countries Sweden, Norway and Denmark score the highest. The exception to this trend is Finland, which for the overall composite and the three dimensions of participatory engagement ranks in the middle of the table. In the domain of Values, however, Finland is ranked 3rd. The group of Scandinavian Countries is followed by Central European Countries: Among them, the highest score is recorded by Belgium, followed by Austria and Netherlands, Luxembourg and Germany. The group of Anglo-Saxon countries plus Finland are ranked from the 9th to the 11th position and they perform much better than France, Mediterranean countries and Slovenia. Finally, in general, it is Eastern Europe and Greece that figure in the lower end of the ranking.

The results among the different dimensions are shown in Table 2. In general, Nordic Countries (especially Sweden) show top performances in all the different dimensions, presenting a valuable consistency in their performances. In contrast, Central European Countries show performances with different profiles; whereas the Netherlands and Luxembourg have consistent performances in all dimensions considered, Belgium compensates for low scores in the dimension of Values with outstanding performance in Political Life.

Table 2 – *The Ranking of the Active Citizenship Composite Indicator.*

Rank	Country	Score (median)
1	Sweden	1.017
2	Norway	0.731
3	Denmark	0.600
4	Belgium	0.565
5	Austria	0.436
6	Luxembourg	0.324
7	Netherlands	0.312
8	Germany	0.295
9	Ireland	0.121
10	Finland	0.056
11	United Kingdom	−0.018
12	France	−0.286
13	Spain	−0.352
14	Italy	−0.470
15	Slovenia	−0.474
16	Portugal	−0.565
17	Greece	−0.789
18	Poland	−0.806
19	Hungary	−0.833

Moreover, looking at the individual indicator included in the dimension of Protest and Social Change (Civil Society), the Nordic countries, where NGOs thrive, have high scores, and they are followed by Western European countries. The lower-scoring countries are from Eastern and Southern Europe. The driver of this result is mainly the sub-dimension of protest which is relatively high for all countries considered, whereas the Achilles heel is participation (especially in trades union). The low score of Poland and Hungary is especially driven by a low score for in volunteering working in organisations (6.5% for Poland and 3% for Hungary, compared with the 30% of the top performer) and in participation in human rights organisations (1% for both countries, while the top performer reaches 4.3%).

Portugal shows better performance in this latter variable (2%) and Greece is particularly strong in the dimension of protest.

The dimension of Community Life shows a slightly different picture. Here high scores are achieved by Belgium and the UK as well as by the Nordic countries. Participation and membership in sports and cultural activities are the driving force of the result. The low position of Italy is mainly the result of low participation and voluntary work and Spain compensates for its low score in participation and membership with high scores for parent–teacher organisations. For Southern Europe, the variable non-organised help is probably not sufficient to represent the informal networks and family support that characterise this region. In countries like Italy, for example, activities like preserving the food heritage (e.g. the Slowfood movement), or keeping cities lively with evening street activities could be considered relevant. Community participation scores low in Eastern Europe, especially in Poland. Furthermore, in Poland religious activities are more frequent than elsewhere in Europe. The dimension of Democratic Values shows a significantly different pattern from the previous dimensions, with some countries demonstrating quite different behaviour and overall fewer regional distinctions. Poland scores quite well in this index and enters the top five. In contrast to the other dimensions, Portugal also scores well in eighth place. In addition, Finland and Luxembourg join Sweden on the top three. The position of Belgium results from its relatively lower scores in the indicators on values on human rights as only about 2/3 of Belgian respondents said that they would give the same rights to immigrants and about the same number considered important the approval of laws against discrimination in the workplace or against racial hatred. In Sweden the proportions were closer to 90% and 80%, respectively.

Finally, in the dimension of Representative Democracy, Austria and Belgium achieve high scores along with the Nordic countries. Austria is ahead of the Nordic countries (in spite of a relatively lower value for women's participation in national parliament), the only occasion in all four dimensions of Active Citizenship that this region does not score the highest. Austria's high score is partly due to the very high number of persons who are involved in political parties. Belgium ranks high in this dimension as a result of its policy of compulsory voting. France and UK perform less well in this dimension than in the previous two indices. Eastern European and some Southern European countries have lower scores. Poland has low voting scores but performs relatively well in donating money to political organisations, whereas Hungary performs well in democratic values and voting (75% in national elections and 38% in European parliament elections) but not in participation in politics. Overall the countries that perform better are not those with the highest voting rates for national or European parliaments but those where participation in politics is higher.

Table 3 – *Ranking of the four pillars of the composite indicator.*

Rank	Country	Protest and social change	Community life	Democratic values	Representative democracy
1	Sweden	2	2	1	2
2	Norway	1	1	4	7
3	Denmark	3	6	7	3
4	Belgium	4	3	18	1
5	Austria	5	9	9	4
6	Luxembourg	11	10	2	5
7	Netherlands	6	5	11	8
8	Germany	7	7	10	6
9	Ireland	10	8	6	13
10	Finland	12	13	3	9
11	United Kingdom	8	4	13	15
12	France	9	11	16	16
13	Spain	14	14	12	10
14	Italy	15	17	15	11
15	Slovenia	13	12	14	17
16	Portugal	16	15	8	14
17	Greece	18	18	19	12
18	Poland	19	19	5	19
19	Hungary	17	16	17	18

3. Modelling the relation between Active Citizenship and its determinants.

In order to deepen the analysis and provide relations with possible socio-economic and behavioural variables, in this paper, the active citizenship composite indicator is computed at the individual level. Using the individual score of this composite indicator it is possible to study the determinants which foster the level of active citizenship among the individuals. This analysis allows us to understand how the level of Active Citizenship varies with respect to the level of the all variables considered and to identify the drivers of the phenomenon and providing an evidence base for policy development providing an evidence base for policy development. Based on these reasons, the next step of this analysis is to investigate the existence of any multivariate relation between the considered variables and the level of active citizenship; in other words we need to model the relation between active citizenship and its determinants.

3.1 *The Methodology.*

The nature of data in the dataset presents a nested pattern of variability: in particular we have a nested source of variability due to individuals and countries. In literature this type of data are known as hierarchical or nested data and are modelled by using multilevel models. Here we present the best way to deal with multilevel approach by challenging both substantive and statistical motivations.

In general multilevel data structures exists if some units of analysis can be considered as a subset of other units, like for instance time series for different countries, individuals grouped in clusters or in countries. The goal of multilevel is to account for variance in a dependent variable which is measured at the lowest level of analysis by considering information from all levels of analysis: a multilevel data structure may count more than one level of analysis (Snijders and Bosker, 1999). The substantive motivations of using multilevel analysis are different: the first reason is the possibility to combine multiple level of analysis in a single comprehensive model by specifying predictors at different levels: in this way, spanning multiple level of analysis the model suffers less for misspecification than models with single levels. The second reason for using multilevel models is that it is possible to specify cross levels interactions. In this way we can detect if the causal effect of lower level predictors is conditioned by higher level predictors.

In additions to these substantive motivations there are also important statistical motivations for using multilevel models. In particular ignoring the multilevel structure of data carries significant statistical costs in term of possibly incorrect standard errors. In other words if individual levels, for example citizens, are influenced by contextual factors, then individuals sampled by the same context share common behaviors, that is the observations at the individual level are influenced by each other.

In terms of statistical models this mutual influence violates the assumption that the errors are independent. The violation of this assumption produces too low standard errors and consequently the t test tend to be too high, in other words predictors appear to have significant effect when in reality they do not have. Clustering in multilevel data structures pose a challenge to statistical analysis. One approach to solve this problem is to absorb contextual and subgroup differences by using dummy variables but this practice even if it is able to take into account the subgroup effect, is not able to explain why there is an effect at the subgroup level; dummies are not able to explain cross level interactions.

The best way to analyze hierarchical data is by using multilevel models which provide correct estimations of standard errors and allows simultaneous modeling of individual level and country level effects. We performed our analysis with Stata software

3.2 *Model selection.*

The case study we deal with has a structure which presents a hierarchical structure with two different levels, individuals, at the lower level, and countries at the higher level. The models we performed are presented in the table 4 which shows deviances for each models defined as minus twice the natural logarithm of the likelihood.

Table 4 – *Model Selection based on deviance test.*

Model		-2Loglikelihood	Δ– -2Loglikelihood	Δ df
0	Intercept	11292.5044		
1	0 + random variation at country level	7858.6448	3433.8596	1
2	1 + Individual variables	4386.733	3471.9118	2
3	2 + country characteristics	4363.4656	23.2674	20

The deviance can be regarded as a measure of lack of fit between model and data, as we can see from the table 4 we interpret the deviance as values differences for the four models we run. The deviance difference follows a $\chi 2$ distribution with degrees of freedom equal to the number of parameters to be estimated. The first model we run is the null model which includes only the intercept and allows variation only at individual level. Model one is a two levels model and the intercept varies across individuals as well as across countries. By confronting the two models we can conclude that the second one is better than the first one because there is a large improvement in the deviance. This means that the level of active citizenship significantly varies both at individual and countries level. The difference between the two deviances is 3434 and it is significant with one degree of freedom. We can calculate the intraclass correlation coefficient ρ as proportion of variance that is accounted for the group level: in model 1 ρ=0.016 which is high, compared to similar case study related to social context. This means that there are significant similarities between individuals in the same country and the use of hierarchical models is then justified. Since we are interested in characterizing the individual identikit of active citizens we introduced variables at the individual level in the model, which, as we can see from table 4, improve significantly the model: the deviance decrease of 3472 with two degree of freedom and the variance at individual level is decreased significantly, from 0.084 to 0.075, as we can see from

table 42. In this model we assume that countries specific regression lines are parallel, this assumption allows individual varying differently across countries, but countries differ with respect to the average value of the dependent variable. In model 3 we introduce the country variables because we want to define the peculiarity of each country taking into account the social, economic and cultural dimension. As we can see from table 4 the model improves significantly, a change of 23 in the deviance with 20 degree of freedom. By introducing group level variables the unexplained variance at group level decreased from 0.01 to 0.001, while the variance at individual level is unchanged, this means that the model catches the group level effect.

3.3 *The model.*

In this section we present the model selected according with the procedure introduced in the previous paragraph. The model has been performed on a set of 14 European Countries, which are almost all the old member states plus Norway. The total number of observations considered in the model is equal to 24915. In particular the countries included in the analysis are:

Table 5 – *List of countries included in the analysis.*

Austria	Finland
Belgium	United Kingdom
Germany	Greece
Denmark	Italy
Spain	Luxembourg
Netherlands	Norway
Portugal	Sweden

The remaining countries (Poland, France, Hungary, Slovenia and Ireland) have been excluded from the analysis due to the fact that some individual level variables were missing. People in education has been excluded from the analysis so, the results are referred to those who have already completed their formal education.

We performed a linear random slope model and the set of individual variables included in the model is listed in the following table.

Table 6 – *List of Individual Variables included in the model.*

Age	age of the respondent at the time of the interview.
Gender	dichotomous (male=1, reference category)
Years of education	self reported number of years of formal education completed
Lifelong Learning	participation at conferences, courses or other learning activities during the past 12 months (yes/no=reference category)
Attendance of Religious Services	Attendance of Religious service apart special occasion: (1: Never, …, 6:Every Day) – recoded with inverted scale
Reliousness	How religious are you: subjective feeling (0-10)
Watching TV	average hours spent in watching TV on a weekday (0:never 7: more than 3 hours)
Listening to the radio	average hours spent in listening to the radio on a weekday (0: Never, …, 7:More than three hours)
Reading Newspaper	average hours spent in reading newspaper on a weekday (0:never 7: more than 3 hours)
Domicile	urban=0/rural=1
Self Reported Income	self reported income of the respondent, coded following the ESS coding (from 1 to 12)
Main Activity	- Main Activity: our elaboration from the original ESS question (with: 1-employed: in a paid work/ military service ; 2-unemployed: unemployed, looking for a job; 3- Retired: retired; 4- Other: Sick, Housework, Other.) - Recoded in 4 dichotomous, mutually exclusive variables

To facilitate the coefficients comparison all the variables have been standardized using the z-score formula. During the analysis the quadratic effect of some variables has been included in the model.

Then, at the country level the variables considered in to the model are shown in the following table.

Table 7 – *List of Country level variables included in the model.*

GDP pro capita	year 2002, Eurostat source
GINI index:	year 2002 (2001 or 2003, when 2002 was not available)
Years of Education	computed as country variable
Religious Heterogeneity	Hello index computed on ESS data

Due to the country level variables considered, the individual level variables "years of education" and "self-reported income" have been standardized at the country level in order to avoid the inclusion of redundant information.

The model has been applied to the entire set of countries considered in the analysis, so the model has to be read for the entire Europe. The application of this model to clusters of countries is not possible due to the collinearity problem: not enough countries for the number of country level variables included in the model. Furthermore, we ran a new model to the four clusters (Nordic, Continental, Mediterranean and Anglo-Saxon Countries) with the same set of individual variables and a restricted number of country level variables. The results recorded in the 4 clusters are approximately the same. For this reason, we present in this report only the multilevel model referring to the whole of the dataset (14 European countries). The results of the multilevel models are presented in table 8. Since we are interested in sketching the identikit of active citizens in Europe we present here first the discussion on the effect of the individual variables and then on country level variables.

Age and Active Citizenship

The effect of age on active citizenship is significant and has a negative quadratic effect. This means that the effect of the age is positive until reaching a maximum and then this effect start to decrease. Ceteris paribus for the effect of the other variables, effect of age recorded a maximum for people of 58 years old, after this level the effect of age start to decrease. Moreover, older people are more active than the young generation. This result follows previous research in the field that through out the lifecycle it is the middle-aged who participate much more. It equally points towards the downwards trend in participation levels from the Baby Boomers/ '68 generation who have always been active in comparison with the new generation of less engaged youth

Gender and Active Citizenship

As shown in table 8 the gender is not significant: no statistical difference is found for the level of active citizenship between male and female, this means that the level of active citizenship is not influenced by the gender.

Education, Life Long Learning and Active Citizenship

As anticipated from the previous literature, the effect of education is strongly positive and is strengthened by considering its quadratic trend, which is positive and reinforces the effect of the variable. Ceteris paribus, the level of active citizenship increase when the number years of education completed increases. As this effect is quadratic, people with a great number of years of education participate in much more active citizenship activities than the others.

Table 8 – *Results of the multilevel analysis.*

Dependent: Individual Active Citizenship		model 0		model 1		model 2		model 3	
	Fixed effect	Coef.	P>|z|	Coef.	P>|z|	Coef.	P>|z|	Coef.	P>|z|
Individual Level	**age**					**0.021**	0.000	**0.021**	0.000
	age (quadratic)					**-0.018**	0.000	**-0.018**	0.000
	gender:male					0.002	0.650	0.002	0.632
	years of education					**0.064**	0.000	**0.064**	0.000
	years of education (quadratic)					**0.004**	0.021	**0.004**	0.022
	LifeLongLearning					**0.092**	0.000	**0.092**	0.000
	Attendance Relig. Service					**0.027**	0.000	**0.027**	0.000
	Religious Feeling					**0.024**	0.000	**0.024**	0.000
	Religious Feeling (quadratic					**0.015**	0.000	**0.015**	0.000
	Citizenship					0.015	0.215	0.015	0.203
	Watching Tv					**-0.021**	0.000	**-0.020**	0.000
	Watching TV (quadratic)					**-0.008**	0.001	**-0.008**	0.001
	Listening to the Radio					0.001	0.522	0.001	0.526
	reading Newspaper					**0.032**	0.000	**0.032**	0.000
	Domicile:Rural					**0.020**	0.000	**0.020**	0.000
	Self Reported income					**0.013**	0.000	**0.013**	0.000
	Main Activity: unemployed					-0.010	0.118	-0.010	0.118
	Main Activity: retired					0.011	0.176	0.011	0.183
	Main Activity: other					**0.076**	0.000	**0.075**	0.000
Country Level	**GDP pro capita**							**0.001**	0.002
	Gini index							**-0.022**	0.001
	Years of Education (country mean)							-0.041	0.061
	Religious heterogeneity							**0.227**	0.011
	Constant	**0.068**	0.002	**0.068**	0.034	0.02318	0.508	**0.761**	0.044
Random-effects Parameters		Estimate	Std. Err.	Estimate	Std. Err.	Estimate	Std. Err.	Estimate	Std. Err.
Level two random effect									
var(_cons)				0.016435	0.00624	0.01484	0.0056	0.00276	0.00107
Level one variance									
var(Residual)		0.10011	0.001	0.084756	0.00083	0.07533	0.0008	0.07533	0.00081

A fact which has been less investigated in debates on education and its relationship with participation is the relationship between lifelong learning and levels of active citizenship. Lifelong learning has also a considerable positive effect on the level of active citizenship. In fact, people who attended conferences or other learning activities in the past 12 months have a much higher level of active citizenship than those who do not participate in Lifelong learning. This result

confirm the primary role which education has in fostering and promoting active citizenship

Religion and Active Citizenship

The effect of Religion on Active Citiztizenshiop have been tested through the inclusion of two variables in the model: Importance of Religion and attendance of religious services apart from special occasions. The two variables show a very interesting picture of the respondent behavior and clearly show the effect of religion in the model. In particular the importance of religion which is measured with a Likert scale 0-10. The quadratic effect have been included in the model too. The effect found is quite unusual and show an U-shape. In fact, ceteris paribus, people declaring that in their life the religion has an importance equal to 0 have a level of active citizenship higher than those who declared an importance of religion varying from 1 to 6. Then, a higher level of active citizenship is recorded by those who declared an importance of religion greater than 6. In some sense a possible interpretation can be that people having clear ideas in their mind about religion (either absolutely no importance or very important) have a higher level of active citizenship with respect to those who are a little more vague about the role of religion in their lives.

The effect of attending religious services is significant, linear and positive. So increasing the frequency of attendance at religious services increases also the level of active citizenship. This result is in-line with the previous one: people who are really religious (religion is very important and they attend religious services) has a higher level of active citizenship with respect to the others.

Citizenship and Active Citizenship

We introduced in the model the legal citizenship variable, however, being a citizen of the country is not significant and has no effect on active citizenship as shown in table 8.

Media Impact on Active Citizenship

In order to assess the effect of the media on active citizenship, we included in the model variables measuring the time spent by the respondent in watching TV, listening to the radio and reading newspaper. Firstly the variable "time spent in watching TV on a average weekday" was included in the model together with its quadratic effect which turned out to have a negative sign. The result is very interesting and the inclusion of a quadratic variable gives a U-shape to the effect of TV permitting a more exhaustive analysis. People who do not watch TV have a lower level of active citizenship than those who watch TV for one hour per day and use the TV to be informed with the news. After that value, increasing the time in

watching TV decreases the level of active citizenship in a very consistent way. TV has a very negative effect for people who watch TV for more than 2 hours per day.

Then the variable "listening to the radio" is not significant in the model and has no effect on the level of active citizenship.

Finally, the variable "reading newspaper" has been found to have a positive effect on active citizenship. Its effect is positive and linear. Increasing the time on reading newspaper increases also the level of active citizenship. Thus certain forms of information gathering have a positive effect on participation whilst watching tv for long periods has negative effect and listening to the radio has no effects.

Domicile and Active Citizenship

As we were interested also to discover if living in cities or in the country side influences the level of active citizenship. The variable we used is an elaboration of the original "domicile of the respondent" which has been recoded in Urban/Rural as a dummy variable. The result is significant and shows that people living in a rural area have a higher level of active citizenship. This results was quite surprising considering those in the countryside have typically further to travel to participate in activities, however, and as noted by Putnam 2001, communities in the countryside are often stronger than in the towns

Self-reported household income and Active Citizenship

The variable measuring the economic aspect of each individual has a significant positive effect and shows that the higher the household income the higher the levels of active citizenship recorded by the respondents. This result is confirmed also by GDP, which has a positive sign. We can interpret both the variables as the level of active citizenship is higher for individuals with high household income and for countries with a high GDP.

Employment and Active Citizenship

We also studied if the different professional status influences the level of active citizenship. The "main activity" variable presents no difference on the effect of active citizenship if the respondent is employed (reference category), unemployed or retired. The only category which turned out to be significant is "others": (housewives, not looking for a job, others). People belonging to this category have a higher level of active citizenship largely we would suspect from having a greater amount of time to participate.

Country Level Variables

Since we are interested also to know the differences between countries in the level of active citizenship we introduced country level variables. The multilevel model we run also allows us to define country characteristics. We introduced four second level variables each for a different dimension which contributes to define the country dynamics like the economic, the social, cultural and religious one. As we can see from table 8 GDP pro capita Gini Index and Religious heterogeneity are significant. The average years of education by country are not significant. The interpretation of these results are that the level of active citizenship is higher in countries with a higher GDP pro capita, a lower GINI index, so a higher level of income equality, and a greater religious heterogeneity.

5. Conclusion.

The results of our research at the individual level predominantly support the trends in the current literature in terms of individual characteristics of age - the young participate less (Putnam 2001), gender - is not significant (Norris 2002), education on an individual level being highly important (Dee 2004, Finkel 2003, Print 2007, Galston 2001, Verba, Schlozsm and Hoskins et al 2008) and income - the more you have, the more you participate (Verba, Slozman and Brady, 1995). Our empirical results also sustain the analysis of Putnam and De Tocqueville concerning the link between religious attendance and active citizenship and the location of the countryside as a stronger bed of community spirit as opposed to the city. In addition, our results also enhance the argument put forward by Putnam that those without occupation and not looking for work, which as a group is dominated by housewives, provide substantial community support in terms of volunteering, participation in associations and generators of social capital (Putnam 2001). Finally our results also support Putnam's thesis on the negative effect of watching television (Putnam 2001). In addition to providing support towards the previous literature results, the empirical analysis in this article has identified a number of new and intriguing findings concerning the individual characteristics of the active citizen, for example, deepening the understanding of religious beliefs. According to our results, active citizens typically have a clear conviction of the importance of religion in their life (either religious or not religious). Thus the persons who are sure that they are not religious are as active as those who are sure that they are. It is the persons who lack a strong belief who are not active. Thus a motivating factor for participation can be considered to be a strong conviction towards religion and not a religious belief in itself.

The second interesting finding is the relationship between active citizenship and lifelong learning. Previous research by Deakin Crick et al., (2005) and Hoskins and

Deakin-Crick (2008) has shown a relationship between citizenship knowledge and values, and the knowledge and values needed for learning providing evidence that education strategies that facilitate one could aid the other. However, to the best of our knowledge, this is the first time that actual participation in lifelong learning and the practice of active citizenship have been identified to be empirically related. Thus active citizens are also active learners and vice versa and that the motivation to participate in society is broader than these individual phenomena and the types of societies and government actions that facilitate one can be considered to be beneficial towards the other.

The country level features that facilitate greater participation in active citizenship are equality, wealth and tolerance towards diversity. In terms of equality the results show that the more equal societies are in terms of distribution of wealth the higher the levels of active citizenship. These findings follow previous research such as Wilkinson and Pickett (2009) that equal societies tend to be more beneficial for most social and health outcomes. The high performing countries in Europe on active citizenship also tend to be the wealthy countries measured by their GDP, in this regard there are two groups of countries: poorer countries that are below the GDP average and have below average participation in active citizenship and more wealthy countries that have higher levels of active citizenship reflecting a two speed Europe. Greater levels of equality also increased average levels of education but unlike years of individual education average levels of education was not found to be associated with active citizenship. In addition to these findings, it is not only equal countries that do well on participation levels of active citizenship it is also the countries that are more tolerant towards other religions who have higher levels of active citizenship measured in terms of religious heterogeneity. This means that in countries with more diversity of religions there are also higher levels of active citizenship. These results are quite the opposite to Huntington's thesis on the clash of civilizations that proposed a lack of social cohesion as a result of greater diversity of religions.

References

Dee, T.S. (2004) "*Are There Civic Returns to Education?"* Journal of Public Economics, 88, 1697-1720.

Delli Carpini, M. and Keeter, S. 1996. *"What Americans know about politics and why it matters."* New Haven: CT:Yale University

Deakin Crick, R., Tew, M., Taylor, M., Durant, K. & Samuel, E. (2005) *"A systematic review of the impact of citizenship education on learning and achievement."* Research evidence in education library. London.

Education Council (2005) "*Council Conclusions of 24 May 2005 on New Indicators in Education and Training, Brussels, 6 October*." Brussels: Education Council.
Education Council (2007) *"Council Conclusions on a Coherent Framework of Indicators and Benchmarks for Monitoring Progress towards the Lisbon Objectives in Education and Training."* Brussels: Council of the European Union.
European Commission (2007e) *"Progress towards the Lisbon Objectives in Education and Training Indicators and Benchmarks."* Staff working document. Luxembourg: Office for the Official Publications of the European Communities
European Commission (2008) *"Progress towards the Common Objectives in Education and Training Indicators and Benchmarks."* Staff working paper. Luxembourg: Office for the Official Publications of the European Communities
Galston, W. (2001). *"Political Knowledge, Political Engagement and Civic Education"*. Annual Review of Political Science, Vol. 4, 217-234.
Hoskins, B. (2006) *"A Framework for the Creation of Indicators on Active Citizenship and Education and Training for Active Citizenship."* Ispra: Joint Research Centre.
Hoskins, B. and Ruth Deakin Crick (2008) *"Competencies for learning to learn and Civic competence: different currencies or two sides of the same coin?"* CRELL research paper, EUR 23360, European Commission: Italy.
Hoskins, B. and Mascherini, M. 2009. *"Measuring Active Citizenship through the Development of a Composite Indicator."* Journal of Social Indicator Research. 90 (3) 459-488.
Hoskins, B., Villalba, E., Van Nijlen, D. & Barber, C. (2008) *"Measuring Civic Competence in Europe: a composite indicator based on IEA Civic Education study 1999 for 14 years old in school."* Ispra: European Commission. EUR 23210 EN.
Hoskins, B., Jesinghaus, J., Mascherini, M., et al (2006) *"Measuring Active Citizenship in Europe."* Ispra: European Commission EUR 22530 EN.
Kahne, J. and Sporte, S. 2008 *"Developing Citizens: The impact of Civic Learning opportunities on students' Commitment of Civic Participation."* American Educational Research Journal. Vol. 45 No.3 pp738-766.
Lauglo, J. and Oia, T. 2008. *"Education and Civic Engagement among Norwegian Youth". Policy Future in Education."* Vol 6 No. 2. pp 203-223.
Marshall, T. (1950). *"Citizenship and social class and other essays."* Cambridge: Cambridge University Press.
Mascherini, M., & Hoskins, B. (2008). *'Retrieving expert opinion on weights for the Active Citizenship Composite Indicator'*, European Commission – Institute for the protection and security of the citizen – EUR JRC46303 EN.
Moldan, B., & Billharz, S. (1997). *"Sustainability indicators: Report of the Project on Indicators of Sustainable Development"*. SCOPE 58. Chichester: John Wiley & Sons.
Huntington , S (1996) *"The clash of civilisations: remaking of the new world order."* New York, Simon & Schuster.
Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Hoffman, A., & Giovannini, E. (2005). *"Handbook on constructing composite indicators: Methodology and user guide"*. OECD Statistics Working Papers 2005/3, OECD Statistics Directorate.

Norris, P. (2002) *"Democratic Phoenix: Reinventing Political Activitism."* Cambridge, Cambridge University Press.
Snijders, T. and Bosker, R. (1999), *"Multilevel Analysis"*, London, Sage publication
Verba, S., Schlozman, K. and Brady, H. 1995. *"Voice and equality: Civic voluntarism in American Politics."* London: Harvard University press.
Verba, S., & Nie, H. (1972). Participation in America: Political democracy and social equality. New York: Harper and Row.
Wilkinson, R., Pickett, K., 2009, *"The Spirit Level, Why More Equal Societies Almost Always Do Better,"* Allen Lane Publisher.

SUMMARY

Facilitating Active citizenship is one of the European Commission's strategies for increasing social cohesion and reducing the democratic deficit across Europe within the context of the wider Lisbon process. In this context, this paper provides an evidence base for policy development, identifying the socio-demographic characteristics and determinants of active citizens and those who for one reason or another participate much less. The report provides a detailed identikit of the active citizen from 2002 across 14 European countries Austrian, Belgium, Germany, Denmark, Spain, Finland, United Kingdom, Greece, Italy, Luxemburg, Netherlands, Norway, Portugal, Sweden (the complete dataset available for this research is only available for the majority of old member states of the European Union and European Economic Area). The results of our analysis, based on a multilevel regression model, provide a clear identikit of the active citizen in Europe and the drivers of the phenomenon are identified both at the individual and at the country level. The picture provided is quite interesting and shows that the level of Active Citizenship is higher in countries with a higher level of GDP with a more equal distribution of income and a more heterogeneous religious climate. Moreover, at the individual level, the strongest determinant of active citizenship is education and participation in lifelong learning activities which can permit some action to policymaker in order to foster the participation in civil society of the the new generations which quite passively do not take part in the democratic life of our societies.

Massimiliano MASCHERINI, Ph.D., *Scientic officer – Joint Research Centre of the European Commission*, Ispra (VA).
Andrea SALTELLI, Ph.D., Head of the Econometric and Applied Statistics Unit – *Joint Research Centre of the European Commission*, Ispra (VA).

Rivista Italiana di Economia Demografia e Statistica Volume LXIII nn. 1-2 – Gennaio-Giugno 2009

MEASURING LONGITUDINAL POVERTY

Daria Mendola, Annalisa Busetta, Anna Maria Milito

1. Introduction

Recent studies show that, in some countries, there are good reasons to believe that experiencing poverty in any given time period may not represent a severe disadvantage if it is unlikely to ever happen again (see, among others, Mendola *et al.* 2009). In contrast, an extended spell of time spent below the poverty line will be considered, for most individuals, as an economic disadvantage and drive them to social exclusion. In this paper we argue that an appropriate measurement for economic disadvantage can be provided by constructing a measure of poverty across time.

In fact, analysing the determinants behind why some households or individuals can escape poverty, while others cannot, requires a dynamic approach based on longitudinal data. Similarly, analysis of poverty entry, stability of households' income situation, and whether poverty is a repeated phenomenon or not, all require a dynamic perspective. The dynamic aspects are consequently the means within which poverty occurs and shapes the experience of being poor (Walker and Ashworth, 1994; Muffels *et al.*, 2000). So it is obvious that the time dimension is of crucial importance. Strategies to cope with short-run poverty are completely different from the ones to face long run poverty, whereas the reduction of expenditures is untenable over a long lasting period of deprivation.

On the other hand, at macro level, if income mobility is high and poverty is experienced for a short period by a larger proportion of the population, then the probability of being poor is equally shared than in the situation of very little income mobility (Fouarge and Layte, 2005).

This paper addresses the issue of measuring longitudinal poverty proposing a path-dependent poverty index which allows to measure the magnitude of poverty persistency. It is structured as follows. Firstly, we discuss the issue of measurement of poverty across time, which forms the conceptual framework of our analysis, undertaking a brief review of the literature concerning the longitudinal of poverty (§ 2). Secondly we introduce our definition of individual longitudinal poverty, and propose some ideal properties which a good index of longitudinal poverty should have (§ 3). Thirdly we go deep presenting our longitudinal poverty index and its main features (§ 4). Then, construct, content and criterion validity of the index is

proved by analyses based on a panel of European young people from European Community Household Panel (§ 5). Finally, some remarks on the further potentialities of the index.

2. Studying poverty in longitudinal perspective

In recent years, due to the increased availability of longitudinal surveys, a wide literature on the dynamics of poverty has developed. This literature demonstrates that: a. " longer time spent in poverty increases the probability of being poor in the future", b. " longer time spent in poverty drives individuals to worse and worse situations", c. "increased time in poverty is associated with a wide range of detrimental outcomes", and that d. "being longer in poverty reduces the chances of being able to climb out of poverty".

As expected, there exist different measures and methodologies to study poverty across the time. Before to start a brief review of the approaches used to define and measure poverty over time, it is worthwhile to highlight that there is a terminological heterogeneity in defining poverty over time. Chronic poverty (e.g. Foster 2007; Bossert *et al.* 2008; Calvo and Dercon 2007; Jalan and Ravallion, 1998), lifetime poverty (e.g. Hoy and Zheng 2008), permanent/persistent poverty (e.g. Mendola *et al.*, 2009; Gaiha and Deolaiker, 1993), intertemporal poverty (Porter and Quinn, 2008), recurrent poverty (Fouarge *et al.*, 2005), longitudinal poverty (e.g. Muffels *et al.*, 2000), and long-run poverty are only some of the terms that can be find in literature referring to the concept of poverty observed along a reference period. Indeed, as it will be explained in the following, behind these different verbal expressions there is sometimes a different conceptualization of who is to be considered poor in a longitudinal view. However we will use indifferently all these terms, when this is not misleading, even though we prefer "longitudinal poverty" which appears to us more neutral/general.

Traditionally, the studies on poverty over time progresses along two different strands: the modelling approach and the indicators approach, even if there are some contributions that try to integrate both. The former try to put in connection poverty measurement with factors determining deprivation, modelling the functional form of the relationship, estimating exit and enter probabilities in poverty status, or investigating on the characteristics of people experiencing poverty and on the determinants of the phenomenon itself. Nowadays we can find a wide number of studies modelling poverty dynamics on the line of the seminal paper of Bane and Ellwood (1986). Here, we do not dwell on this literature because it is not strictly pertinent with the line of our paper.

The indicators' approach is mainly focused on producing new measurement instruments and in developing analytic properties of indices. The introduction of a set of desirable properties for a poverty indicator is due to Sen (1976) who set the axiomatic approach in cross-sectional measurement of poverty. The best-known measures of poverty over time are Jalan and Ravallion's (2000) poverty measures that build the standard of living measure using the average level of consumption over the entire period. A shortfall of these measures of chronic poverty is that they do not refer to the notion of persistence as it is intuitively associated with the concept of chronic poverty, in fact even one single period of deep poverty may be enough to label a chronic poor.

A recent paper by Mendola *et al.* (2009) using all the waves of ECHP data proposes a classification of young Europeans according to the levels of persistence in the poverty status. These result in a categorization of three groups: *permanent poor*, *socially vulnerable*, and *never poor*; the classifications being based on the sequences of poverty spells, i.e. the way poverty and non-poverty spells alternate one the other.

The paper by Mendola *et al*. (2009) as well as papers by Layte and Whelan (2003), Whelan et al. (2003b), Layte et al. (2003) propose definitions of longitudinal poverty that are simply *ad hoc* solutions merely instrumental for describing the most vulnerable strata of the population, for analysing the determinants of the long-lasting hardship. But above all they are not generalizable nor characterized by properties and axioms.

On the contrary there are some studies specifically dealing with indices of poverty in a longitudinal context which supply an axiomatic framework for the measures proposed (Foster 2007; Bossert *et al.* 2008; Porter and Quinn 2008; Calvo and Dercon 2007; Hoy and Zheng 2008). Even if, as highlighted by Foster and Santos (2006), there are no measure neither axiomatic framework yet that are completely satisfactory.

In more recent years some papers proposed "chronic poverty" definitions and measurements that can account for duration in poverty and take care also to the axiomatic properties. This literature is strongly relevant for us so it deserves a more accurate review.

Both Foster (2007) and Calvo and Dercon (2007) propose measures that "*allow some degree of substitution between well beings in different periods provided they are all below the poverty line, but none across it*".

In particular the measure proposed by Foster (2007) identifies the chronically poor by two distinct cut-offs: income cut-off and duration cut-off. The first one is set on the income space: and it defines poor the one whose income is below the poverty line; the second is related to the percentage of time spent in poverty and defines chronically poor who has an incidence of poverty over time higher than

some fixed level.[1] He suggests also another measure to evaluate poverty that is shorter in duration (i.e. "transient poverty"). These measures weight the depth of poverty experienced by those chronically or transient poor before aggregating over time and individuals. At the aggregated level, Foster (2007) presents a class of chronic poverty measures adjusted to account for the duration of poverty.

Calvo and Dercon (2007) put the attention on three main issues: "*the role of compensation over time (whether poverty spells can be compensated for by non-poverty spells); the issue of the discount rate (whether each spell should be given an equal weight); and the issue of the role of persistence (whether repeated spells should be given a higher weight)*". They propose measures dealing with consecutive spells of poverty but accommodate only for two consecutive periods without looking at the whole sequence of poverty, moreover they allow for compensation but only below the poverty line. A paper by Foster and Santos (2006) cross over this last limitation allowing for *"imperfect substitution both below and across the poverty line"*.

The individual "intertemporal poverty measure", proposed by Bossert *et al.* (2008), pays attention to the length of individual poverty spells by assigning a higher level of poverty to situations where, *ceteris paribus*, poverty is experienced in consecutive rather than separated periods. The aggregation of these intertemporal poverty measures is made by the arithmetic mean of the individual intertemporal poverty indices.

3. Our definition of individual longitudinal poverty

The definition of individual longitudinal poverty, that we propose here, is grounded upon both the total number of years spent in poverty along the observation period, and the sequencing of the poverty and non-poverty episodes. We consider longitudinal poverty as an attribute moving along a *continuum* ranging from "absolutely not poor" to "permanent poor", through different degrees of this hardship. According to our definition an "absolutely not poor" is an individual with no episodes of poverty along the observational period or with at maximum only one year of poverty, while a "permanent poor" is an individual who is poor for all the years we observed him/her. The reason why we consider as longitudinally poor a person even if he/she experiences a year of hardship is that we believe that in a longitudinal view an individual is "not poor" even if a transitory reduction of

[1] Also Gaiha and Deolalikar (1993) classify as chronically poor all families that have incomes below the poverty line in at least five of the nine years of observations.

his/her income makes he/she belong to the poor set, since a single spell of deprivation does not reflect a real change in the level of the living conditions.

In our definition all the trajectories of poverty between the upper and the lower bound introduced above correspond to degrees of poverty rising together with the total number of years spent in poverty and above all with the increase of consecutive spells. Indeed we take into account both individuals staying poor consecutively for a certain number of years and the individuals experiencing poverty and non-poverty spells, giving them different degrees of longitudinal poverty. So, not only the consecutive spells of poverty are considered but in general all the years spent in poverty contribute to measure the longitudinal poverty concept.[2]

We agree with Whelan *et al.* (2003a) that "the time dependent nature of poverty is characterized by four dimensions: 1. the length of the observation period; 2. the extent of recurrent poverty; 3. the length of poverty spell; 4. the volatility and stability of poverty statuses over time".

In particular we believe that good properties for longitudinal index of individual poverty (*micro level longitudinal poverty index*) should allow to take into account all these features simultaneously:

1. given the total number of years in poverty, the index should decrease according to the lengthening of the observation period;
2. given the length of the observation period, the index should rise according to the increase of the total number of years in poverty;
3. given both the length of the panel and the total number of years in poverty, the index should increase if the number of *consecutive* years in poverty rises.

The point raised in (3) implies the 4th:

4. given both the length of the panel and the total number of years in poverty, the index should decrease if there is an high number of transitions in and out of poverty status (*volatility*) over time and on the contrary it is expected to increase if there is a low number of transitions in and out of poverty status (*stability*).

[2] Because our primary attention is focussed on individuals experiencing some degree (different than zero) of poverty, in the following -for the sake of simplicity- we will use the term "longitudinally poor" only for those having a non null degree of longitudinal poverty, and will refer to "not poor" for people never poor or for people poor only once in the period.

4. A new Longitudinal Poverty Index: measuring jointly duration and intensity of poverty

4.1 *The rationale and the methodology*

What we want to present in this paper is an indicator of the long-lasting poverty which reaches its maximum value if and only if the individual is poor along the entire period of observation and its minimum value when he/she is out of poverty for all the time.

Given an individual H, let us suppose to have *n* repeated measurements of his/her poverty status along *n* time periods. These measurements compose a poverty sequence of zeroes and ones, where there is 1 when the individual is poor at time *k* (whatever the criterion), and 0 otherwise. Given the length of the observation-period (panel) *n*, it is theoretically possible to have 2^n possible different sequences. Thus if H is always poor in all the *n* waves, his/her poverty sequence is something like (11111….1111), while, if H is always not poor, it results that its sequence is (00000……0000).

Let us suppose we have the observations on the poverty status of three persons for seven waves, let us say H_1, H_2, H_3, so that their observed sequences are respectively $\mathbf{s_{H1}}$= (0000111), $\mathbf{s_{H2}}$=(1100100) and $\mathbf{s_{H3}}$=(1000011). These three individuals have the same poverty hit rate: they all spent three years in poverty during the observed seven years. But, with respect to the persistence in poverty, it is easy to note that H_1 experienced the highest persistence because it underwent three *consecutive* years of poverty, while H_2 and H_3 alternated good and bad times, with shorter spells in poverty, which is typical of socially vulnerable people.

Each sequence can be indexed by *i*. So, for example, in the vector $\mathbf{s_{H1}}$ the first observation is in position *i*=1, the second stays in position *i*=2, and so on (corresponding respectively to first wave of the panel, second wave and so on):

Positions (i)	*1*	*2*	*3*	*4*	*5*	*6*	*7*
Status	0	0	0	0	1	1	1

For each poverty sequence we can consider the correspondence between the observed vector **s** whose elements can be only 1 and 0, and the vector of their positions $\mathbf{s_{pos}}$ =(1234567) whose elements are all the integers from *1* to *n*. Let say that $\mathbf{s}^*$ is the subset of the positions vector $\mathbf{s_{pos}}$ which refers only to integers corresponding to years spent in poverty. For individual H_1 is $\mathbf{s}^*$=(567). Note that $s^* \subseteq s_{pos}$, the equality stands only when the sequence includes only 1s.

Our Longitudinal Poverty Index (LPI) for individual H is:

$$LPI_H = \sum_{i,j \in s^*} (i-j)^{-1} \qquad \text{with i>j} \qquad (1)$$

so that *i* and *j* are the positions in which we can find a 1 in the sequence, and *i* is temporally subsequent to *j*.

In order to compute the index, consider that in $\mathbf{s_{H1}}$ we have three ordered couples (i, j) made with two 1s, in which *i* follows *j*. They correspond to the positions indexed by: (7,6), (7,5), and (6,5). So according to formula (1) we have:

$$LPI_{H_1} = (7-6)^{-1} + (7-5)^{-1} + (6-5)^{-1} = 2.5$$

Following the same procedure as above, from $\mathbf{s_{H2}}$=(1100100) once more we have three couples, (5,2), (5,1), and (2,1). So that:

$$LPI_{H_2} = (5-2)^{-1} + (5-1)^{-1} + (2-1)^{-1} = 1.6$$

and for $\mathbf{s_{H3}}$=(1000011), LPI_{H3} =1.4.

4.2 *Analytical and empirical properties*

To test the ability of the index in measuring the severity of the phenomenon, and in order to distinguish different patterns in the individual poverty life experience here we shall show some results on the analytical property of LPI. Specifically we looked at what happens when the total number of years spent in poverty (*t* in the following) varies, and how the index is influenced by the position of the years in poverty along the sequence.

It results that the index has the following properties (see Mendola and Milito, 2008):

P1. It can be easily proved that, given *n*, the index does not depend merely on *t*, but it is affected by the position of the 1s in the sequence, and principally from the relative distance between each couple of 1s in the sequence. That is, our index decreases in value when, *ceteris paribus*, the distance between the positions of two 1s increases; *vice versa* when the distance between two years of poverty decreases, the index increases accordingly (*path dependence*).

P2. Given the mutual distances between all the couples of 1s, LPI is invariant for the positions of these inside the sequences (*shifting invariance*). This means that to the sequence (000100010) corresponds the same value in the index such as in (001000100) or in (100010000). This is because distances (*i-j*) between the positions of the 1s are equal to three in all the cases shown.

P3. The computation of the index exploits all the years spent in poverty (such as in poverty hit rate) but it gives them a different relevance according to the reciprocal of their distance. That is two or more consecutive years of poverty give a contribution to the index higher rather than two years of poverty interspersed with one or more years of non-poverty (*cumulative hardship*).

P4. (Corollary of P3) When t is constant, LPI increases together with the length of each sub-sequence of consecutive 1s.

P5. All that occurs outside the interval between the first and the last 1s observed does not change the value of the index (*main pattern focus*).

P6. The index ranges from 0 to $\sum_{k=1}^{n-1} \frac{k}{n-k}$, and both the index and its maximum increase with n. Having a maximum allows us to evaluate the magnitude of the phenomenon of longitudinal poverty (*upper and lower boundedness*).

In addition as it is explained in the following section LPI has also some interesting empirical properties and in particular:

P7. LPI and average poverty gap (across the individual sequence) are highly positively correlated. This is an indirect proof that even if LPI does not account explicitly for intensity of poverty it encompasses this information.

Just for the sake of clarity, Table 1 shows the maximum number of different poverty sequences for a panel from 2 to 20 waves and the maximum of LPI.

Table 1 – *Maximum number of different poverty sequences for panel of length n and values of the maximum of LPI.*

N	Number of possible sequences	Maximum value of LPI
2	2	1
3	8	2.5
4	16	4.3
5	32	6.4
6	64	8.7
7	128	11.2
8	256	13.7
9	512	16.5
10	1,024	19.3
11	2,048	22.2
12	4,096	25.2
13	8,192	28.3
14	16,384	31.5
15	32,768	34.8
…		
20	1,048,576	54.8

In conclusion these are good properties, and they are consistent with the concept we want to catch since persisting in poverty consists in keeping on the same status year after year. Accordingly we welcome the fact that our index decreases when the sequence of years in poverty is interspersed with some years out of poverty.

4.3 *The normalised Longitudinal Poverty Index*

The existence of a maximum for LPI allows to build up a normalised index which is a measure more readable. The normalised version of the index, namely LPI*, is provided by:

$$LPI_H^* = \frac{\sum_{i,j \in s^*} (i-j)^{-1}}{\sum_{k=1}^{n-1} \frac{k}{n-k}} \qquad \text{with i>j} \qquad (2)$$

LPI* spans in [0,1], it keeps all the properties from P1 to P5 listed for LPI, and in addition

P8. The longer the reference period the less the relevance of the number of years spent in poverty. That is LPI* is able to distinguish among different degrees of longitudinal poverty (*continuity*). For example considering the following sequences:

n=8	t= 5	(0 **1 1 1 0 0 1 1**)
n=15	t=5	(0 0 0 0 0 0 **1 1 1 0 0 1 1** 0 0)
n= 20	t=5	(**1 1 1 0 0 1 1** 0 0 0 0 0 0 0 0 0 0 0 0 0 0)

we have that LPI* equals respectively 0.36, 0.14, and 0.09[3] while LPI produces the same results (LPI=4.932).

Note that most of the traditional axioms, which are popular in the literature of longitudinal poverty indices, does not apply to LPI and neither to LPI*, while some are hardly translatable in our conceptual framework. This could happen mainly because the axioms refer to poverty over a population, while LPI* is a micro-level index. Moreover often the axioms refer to a hypothetical redistribution of incomes among individuals, while in general income transfers do not affect LPI* with an

[3] Even if in theory the normalized index allows for comparisons among individuals observed for periods of different length, we believe that these should be avoided especially considering the well-known problems deriving from censoring in panel data which are particularly relevant due to the rationale behind the index construction.

exception when the transfer produces a cross over of the poverty line (downward or upward) or a change of status (from not poor to poor or *viceversa*).

Anyway referring to the axiomatic approach to "chronic poverty" indices in Foster (2007), it is easy to prove that LPI* satisfies two upon the three axioms Foster set for the time dimension of an index of poverty. That are *time focus axiom* and *time monotonicity axiom:*

- *time focus axiom* asserts that "*an increase in income during a period when a chronically poor person is not in poverty will not serve to lower chronic poverty at all"*.

In our case, substituting "chronically poor" with our definition of "longitudinally poor", we can say that LPI* satisfies the time focus axiom; in fact our measure ignores the current level of income of a not poor person, since this not affect neither his/her status (0 or 1) nor the positions of this spell inside the poverty sequence;[4]

- *time monotonicity axiom*, according to Foster, says that *"during a period in which a chronically poor person happens to be having a spell outside of poverty, if the income level falls below the poverty line (thus raising the number of duration of poverty experienced by this person), then poverty should rise"*.

This axiom still stands for our index; in fact this is the situation when, given the length of the panel, the total number of years in poverty increases. Note that this axiom is very similar to property P4. given above.[5]

The third time-axiom in Foster scheme (*time anonymity axiom*) need a special attention, while it is not clear what is to be defined anonymity in a longitudinal contest. According to Bossert *et al.* (2008) we believe that "*the negative effects of being in poverty are cumulative, hence a two-period poverty spell is much harder to handle than two one-period spells that are interrupted by one (or more) periods out of poverty"* (see our cumulative hardship property P3). It implies that the property of *time anonymity "under which the sequencing of incomes in individual intertemporal profiles does not affect poverty"* (Foster, 2007) is expressly violated because it represents for us a too severe simplification.

[4] Indeed LPI* is not sensitive also to a decrease in income during a period when an individual is poor.

[5] In the following paragraph 5 we will show that the correlation between the LPI* and the total number of years spent in poverty by each person is 0.934.

5. Assessing validity of LPI*

There are in the literature many domains to evaluate validity of an indicator, but we agree with Schmitz (1993) when she says that many researchers have an unclear understanding about what it means to have a valid indicator. "*There are no agreed-upon procedures for assuring that the indicators measure, or represent, that which we assume they measure*". This lack of agreement is not limited to researchers but extends to developers of indicator systems. "*The validity of current indicators might be less in question had they been developed systematically, beginning with careful definition of an underlying construct [...] and proceeding to the development and testing of standardized measures*".

In order to assess the validity of our index of longitudinal poverty, we refer to the concepts of "construct validity", "content validity" and "criterion validity" (Carmines and Zeller, 1979).

A) "Construct validity" measures the extent to which one purported measure of poverty is correlated with other measures with which a high correlation would be expected.

B) "Content validity" addresses the extent to which the content of a measure is consistent with professional knowledge about longitudinal poverty.

C) "Criterion validity" refers to the correlation of our measure with a gold standard measure, whose validity has been assessed before.

To verify the validity of the proposed index we used a balanced panel data on young European people (aged 16-29 at first wave) from the European Community Household Panel (ECHP). The dataset provides harmonized information on European countries along eight waves from 1994 to 2001. As usual in the year *t* of the interview the income of the past year (*t-1*) is recorded together with the current information on occupation, health, education, marital status, family size, and so on. We made a time re-alignment of income and personal and household characteristics so that we have contemporary information for seven waves only.

A) A way to test the "construct validity" could be to compare LPI* with the total number of years spent in poverty and with the length of the longest spell in poverty for each individual (named *maxspell* in the following). In fact both these measures are supposed to be highly positively correlated with whatever index of longitudinal poverty (in our case with LPI*). Considering only the individuals who are longitudinally poor (LPI*>0) and referring to the sample of young European described above, it results a quite strong relationship: the correlation with the total number of years spent in poverty by each individual is 0.934 while the correlation between the LPI* and the *maxspell* is 0.948.

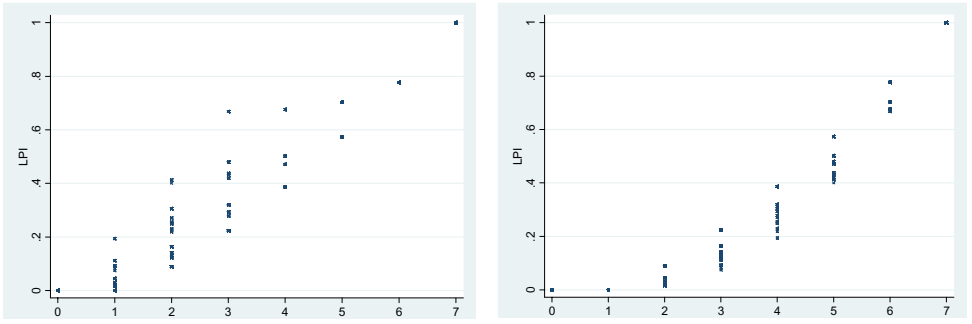
In Figure 1.a, we plot the values of normalised longitudinal poverty index against the *maxspell*, while in Figure 1.b we show the relationship between LPI*

and the total number of years spent in poverty. In both cases, a quite strong relationship between LPI* and the two measures of poverty persistence is evident. But note that:

1) the panel survey that we are using to assess validity is quite short (seven waves) and so, given a consecutive spell of poverty of length four or five, the number of observed different sequences is strongly limited;[6]
2) even with short panel data, such as ECHP, our index is still able to discriminate individuals experiencing different poverty profiles more effectively than *maxspell* and the poverty hit rate (which, with our balanced panel, is proportional to the total number of years spent in poverty) are able to do.

Figure 1 – *Relationship between longitudinal poverty index (LPI*) and other measures of poverty permanence (among young European).*

a. the longest spell of poverty (*maxspell*) b. total number of years in poverty



B) The "content validity" asks LPI* to produce consistent results in the interpretation of the long-run poverty, with an acceptable agreement with qualified literature. In order to prove that our index verifies this property the following box-plots show the distribution of LPI* for some factors usually associated with the poverty status and the most severe forms of deprivation. All the box-plots below refers to the values of the longitudinal poverty index among young European people who have LPI* greater than zero.[7] This choice is in order to make the

[6] We are going to test this issue with a longer panel such as BHPS or SOEP.

[7] As a general note, mind that all the graphs in the following explore only bivariate relationships (i.e. those between LPI* values and some factors traditionally studied in connection with poverty). This means that we do not control for any of the other variables which could interfere on the relationship. So for example in the following we will see that women are more disadvantaged than men, experiencing higher level of longitudinal poverty.

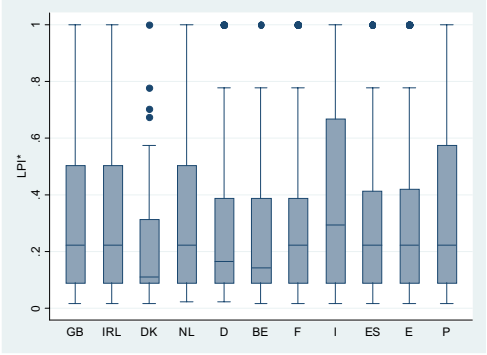
graphs more readable, given that around 70 per cent of the individuals in the sample has at maximum one year of poverty along the reference period.

LPI* shows the expected behaviour when compared with unemployment[8] (Figure 2) and gender effects (see Figure 3 and 4).

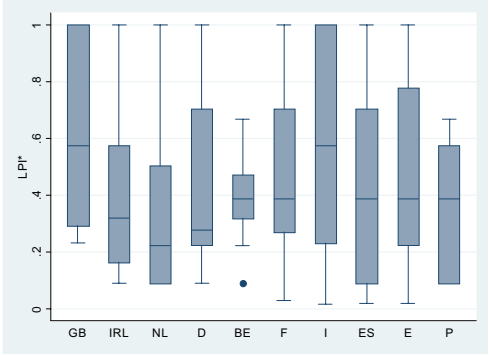
Figure 2 shows, as expected, that being for a long time out of the labour market (because you are a worker or a student) is associated with higher levels of longitudinal poverty.

Figure 2 – *Unemployment and longitudinal poverty among young people in European countries.*

People staying less than 50% of the period unemployed (because worker or student)



People staying more than 50% of the period unemployed

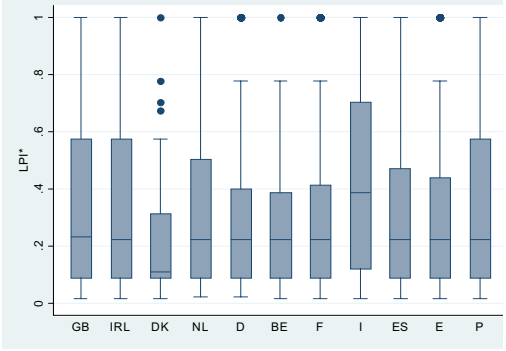


* Graph for values of LPI*>0

A comparison among European countries (Figure 3) shows, as well known in the literature, that not all the welfare systems are equally able to reduce disparities among young people, and to smooth long-run poverty. In particular in Italy the variability in the degrees of individual longitudinal poverty is wider and the first quartile and especially the median level of longitudinal poverty are the highest.

[8] This is a variable accounting for the percentage of the reference period an individual stayed unemployed (more versus less than 50%) (for more details see Mendola *et al.*, 2009).

Figure 3 – *Longitudinally poor young people in European countries.*

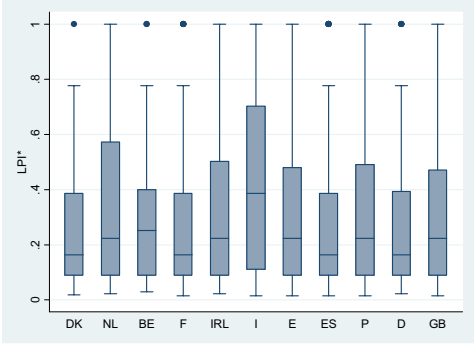


* Graph for values of LPI*>0

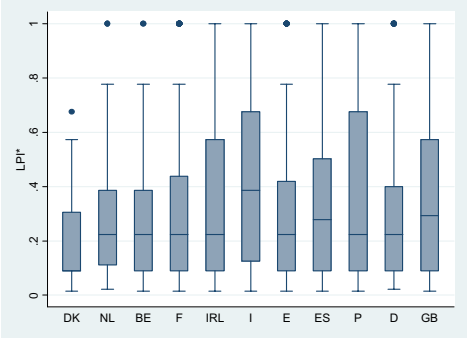
Looking at the distributions of LPI* for men and women we notice that the median value of the longitudinal poverty index is higher for women than for men(not controlling for other factors). Figure 4 gives a comprehensive view of this phenomenon in the eleven countries of the ECHP. Moreover women present a wider inter-quartile range in Italy, Portugal, Ireland and Great Britain while men only in Italy and Netherlands. Some welfare regimes have a strong effectiveness in smoothing out inequalities (see for example Aassve *et al.* 2005; Ferrera, 1996; Fouarge and Layte, 2005). It is the case of Social Democratic countries where the ample attention of welfare system devoted to protect young women show its effect in panel *b* of Figure 4, where women have substantially lower dispersion (inter-quartile range) around the median than men.

Figure 4 – *Longitudinally poor in European countries by gender.*

a. Men



b. Women

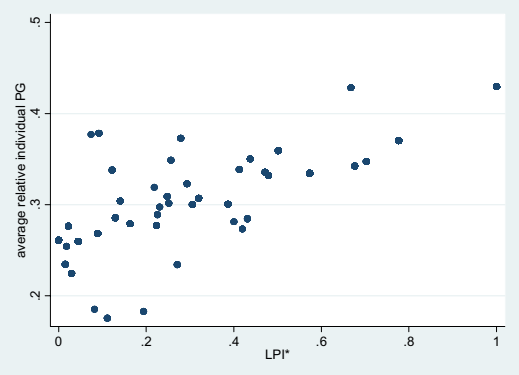


* Graph for values of LPI*>0

C) The "criterion validity" presumes a high correlation with some validated measures of the same concept. Whereas poverty at any given period may not be very informative about the severity of the individual's poverty, its sequencing over time will. That is to say, an individual recorded with several poverty spells is also more likely to be well below the poverty line compared to someone who only experienced poverty for a short period. In other words, our measure of longitudinal poverty is supposed to be also a measure of how severe poverty is.

In order to verify this assumption we computed the correlation among our LPI* and the conditionals means of individual poverty gaps for each of the values of LPI* across the seven waves of the ECHP panel[9]: correlation equals to 0.936. Figure 5 shows how higher persistence in poverty (i.e. higher levels of LPI*) is related to higher values of average individual relative poverty gap.

Figure 5 – *Relation between LPI* and average individual relative poverty gap among young European.*



So at the end the index proved to be a quite valid measure of the phenomenon of poverty across time. In particular it showed the expected performances when put in connection with other easy, but accredited measures of longitudinal poverty. As a next step in the validation procedure we will reserve to put our index in relation with other measures of longitudinal poverty in order to assess the consistency with some longitudinal poverty indices known in literature (among others Calvo and Dercon, 2007; Foster, 2007; Bossert *et al.*, 2008), and to verify the performances of LPI* in longer panel datasets.

[9] As mentioned in theory LPI* is a continuous measure in [0,1], but in this case, due to the small length of ECHP (only seven waves) the index shows only 43 different values.

6. Conclusion

In this paper we moved from the idea that the distinction between temporary and persistent poverty is crucial from a social policy perspective especially with regard to determinants of these two situations. We consider longitudinal poverty as an attribute moving along a *continuum* ranging from "absolutely not poor" to "permanent poor", through different degrees of this hardship. In our definition all the trajectories of poverty between the upper and the lower bound correspond to degrees of poverty rising together with the total number of years spent in poverty and above all with the increase of consecutive spells.

The new index of longitudinal poverty at individual level, that we proposed in this paper, measures the severity of poverty, taking into account the way poverty and non-poverty spells follow one another along individual life courses. The index is normalized and increases with the number of consecutive years in poverty along the sequence. The validation procedures highlighted that the index is strongly correlated with the most common measures of poverty across time (such as the total number of years spent in poverty and the length of the spells of consecutive years in poverty), even if the index is more able to distinguish among people with different poverty profiles. Moreover LPI* produces coherent interpretation of intensity and severity of poverty among young European people. Our longitudinal poverty index could have a vast applicability even outside the field of poverty studies. Indeed it can be applied to all those economic and social phenomena in which it would be interesting to study the persistence in a particular dichotomous status, such as: persistence in an illness status (recovery/recurrence of the illness in a medical follow up); occupational status trajectories (employed/ unemployed properly dichotomised); credit track (pay or not pay the rate of mortgage or other type of loan); persistence of negative (or positive) expectations on the state of the economy (see studies on business sentiment indicators in which there are longitudinal sequences of +/– or growth/fall).

Acknowledgments

This research has been partially supported by founds from two research projects financed by University of Palermo [ORPA051153 and ORPA06CJHS] and coordinated by Anna Maria Milito.

The ECHP data used in the analysis were made available by ECASS (European Centre of Analysis in the Social Sciences) in Colchester (UK), and IRISS at CEPS/INSTEAD in Differdange (LU) during two visiting periods. The visits to ECASS were supported by Access to Research Infrastructure Action under the EU

Improving Human Potential Programme. The visit to IRISS was supported by the European Commission under 6th framework Programme's Research Infrastructures Action (Trans-national Access contract RITA 026040).

We wish to thank all the delegates attending the session of the Annual Conference "Povertà ed esclusione sociale" of the Italian Society of Economics Demography and Statistics (SIEDS) held in Florence in May 2009 for their useful suggestions.

All errors and inconsistencies in the paper are our own.

References

Aassve, A., Iacovou, M., Mencarini, L. (2005). Youth poverty in Europe: what do we know?. ISER working papers n.2. University of Essex, Colchester.

Bane, M.J. and Ellwood, D.T. (1986). Slipping into and out of Poverty: The Dynamics of Spells. *The Journal of Human Resources*, 21(1): 1-23.

Bossert, W., Chakravarty, S. R. and D'Ambrosio, C. (2008). Poverty and Time", Cahiers de recherche 2008-05, Universite de Montreal, Departement de sciences economiques.

Calvo, C. and S. Dercon (2007), "Chronic Poverty and All That: The Measurement of Poverty Over Time," Chronic Poverty Research Centre, Working Paper No. 89.

Carmines, E. G. and R. A. Zeller (1979). Reliability and Validity Assessment. Sage, Beverly Hills, CA.

Ferrera, M., 1996. The "southern model" of welfare in social Europe. Journal of European Social Policy 6 (1), 179-189.

Foster, J.E. (2007). "A Class of Chronic Poverty Measures," Department of Economics, Vanderbilt University, Working Paper No. 07-W01.

Foster, J. E., and Santos, M. A. (2006). "Measuring Chronic Poverty", Mimeo, Vanderbilt.

Fouarge, D. and Layte, R. (2005). Welfare regime and poverty dynamics: the duration and recurrence of poverty spells in Europe. *Journal of Social Policy*, 34 (3): 407-426.

Gaiha, R. and Deolaiker A. B. (1993), "Persistent, expected and innate poverty: estimates for semi arid rural South India". *Cambridge Journal of Economics*, 17, 409-421.

Hoy, M. and Zheng B. (2008). "Measuring Lifetime Poverty," paper presented at the 8th International Meeting of the Society for Social Choice and Welfare, Istanbul.

Jalan, J., and Ravallion, M. (2000). "Is transient poverty different? Evidence for Rural China", *Journal of Development Studies*, Vol. 36, 82- 99.

Jalan, J. and Ravallion M. (1998). "Determinants of Transient and Chronic Poverty, Evidence from rural China", *Policy Research Working Paper 1936*, The World Bank.

Layte, R. and Whelan, C.T. (2003). Moving in and out of Poverty. *European Societies* 5: 167-191.

Layte, R., Whelan, C. and Maitre, B. (2003). Persistent Income Poverty and Deprivation in the European Union: An Analysis of the First Three Waves of the European Community Household Panel. *Journal of Social Policy*, 32: 1-18, Cambridge University Press.

Mendola D., Busetta A., Aassve A. (2009) What keeps young adults in permanent poverty? A comparative analysis using ECHP, *Social Science Research,* 38 (4), 840-857.

Mendola D., Milito A.M. (2008). A new index of poverty persistence, proceedings of IASC 2008: December 5-8, Yokohama, Japan.
Muffels, R.D., Fouarge, D. and Dekker, R. (2000). *Longitudinal poverty and income inequality. A comparative panel study for the Netherlands. Germany and the UK*. OSA-Working paper WP2000-6.
Porter, C. and N.N. Quinn (2008). "Intertemporal Poverty Measurement: Tradeoffs and Policy Options," paper presented at the IARIW 2008 conference in Portoroz.
Sen, A.K. (1976). "Poverty: An Ordinal Approach to Measurement," Econometrica, 44, 219–231.
Schmitz C. C. (1993), Assessing the Validity of Higher Education Indicators, Journal of Higher Education, Vol. 64, 503-521.
Walker, R. and Ashworth, K. (1994). *Poverty Dynamics: Issues And Examples*, Avebury, Aldersho.
Whelan, C.T., Layte, R. and Maitre, B. (2003a). *Poverty, deprivation and time: a comparative analysis of the structuring of disadvantage*. EPAG Working Paper, n.48, Colchester: University of Essex.
Whelan, C.T., Layte, R. and Maître, B. (2003b). Persistent Income Poverty and Deprivation in the European Union. *Journal of Social Policy,* 32(1): 1-18.

SUMMARY

Traditional measures of poverty persistence, such as "poverty rate" or the "persistent-risk-of-poverty rate", do not devote enough attention to the sequence of poverty spells. In particular, they do not put enough attention in underlining the different effects associated with occasional single spells of poverty and the consecutive years of poverty. In this paper we propose a new index which measures the severity of poverty in a longitudinal view, taking into account the way poverty and non-poverty spells follow one another along individual life courses. The index is normalized and increases with the number of consecutive years in poverty along the individual poverty profile. The index is supported by a conceptual framework and characterised via properties and axioms. It is validated and tested on a sample drawn from young European adults participating in ECHP survey.

Daria MENDOLA, ricercatrice in Statistica sociale, Dipartimento di Metodi quantitativi per le scienze umane, Università di Palermo, mendola@unipa.it
Annalisa BUSETTA, assegnista di ricerca, Dipartimento di Metodi quantitativi per le scienze umane, Università di Palermo, abusetta@unipa.it
Anna Maria MILITO, professore ordinario di Statistica sociale, Dipartimento di Metodi quantitativi per le scienze umane, Università di Palermo, milito@unipa.it

Rivista Italiana di Economia Demografia e Statistica Volume LXIII nn. 1-2 – Gennaio-Giugno 2009

SOCIO-ECONOMIC CONDITIONS AND HEALTH INEQUALITIES IN ITALY AT THE BEGINNING OF THE XXI CENTURY

Fausta Ongaro, Silvana Salvini

1. Introduction

Continuing a general trend, the Italian population also benefited in 2007 from further improvements in survival. The estimated life expectancy at birth was 78.6 years for men, and more than 84 years for women. There are few countries in the world that show such a level of average life span: in Europe, Italian men were second only with respect to Swedish and Italian women second only with respect to French (Lanzieri, 2008). A general improvement is also observed in the field of health. However, as for survival, elements of heterogeneity exist even for health among this population. Besides the well-known inequalities of age and gender, persistent differences in health depending on socio-economic status are observed, at least at a descriptive level: "*Utilizzando il titolo di studio come indicatore, si osserva come siano sempre le persone con un basso titolo di studio a presentare peggiori condizioni di salute, sia in termini di salute percepita, che di morbosità cronica. In tutte le fasce d'età la quota delle persone che dichiarano di stare male o molto male triplica o raddoppia tra quanti hanno conseguito al massimo la licenza elementare rispetto alle persone con titolo di studio più alto (laureati e diplomati*)…" (ISTAT, 2007a, p.14).

This study intends to analyse more in depth the relationship between health status and socio-economic conditions in light of the most recent data on health conditions in the Italian population. The positive association between health status and socio-economic conditions has been discussed extensively in the literature. People in lower socio-economic groups more often suffer from disease and disability as well as lower self-perceived health, and this is observed even in developed countries with high levels of public health services (Mackenbach, 1992). Part of this association may be attributable to a "selection" effects of health on education level or occupational position, e.g., due to health problems in early childhood affecting school attainment and therefore on related job opportunities. However, these effects have been found to play only a minor role. People in lower socio-economic conditions are more exposed to health hazards in the physical environment (unfavourable material living condition, e.g., housing or working

conditions**)** (Volkers et al., 2007; Costa-Font, 2008), they more often experience psychosocial stressors, and they are more likely to adhere to unhealthy behaviours (such as smoking, inadequate diet, excessive alcohol consumption, and lack of physical exercise) (Giskes et al., 2002). Moreover, they do not always use the existing health facilities in an optimal way (Van der Heiden et al., 2003). Health inequalities thus are principally a problem of unequal distribution of risk factors and health risks affecting mostly lower socio-economic groups (Commission on Social Determinants of Health, 2007).

This work has two aims.

Firstly, we intend to verify to what extent the differences observed at a descriptive level persist after other individual or contextual (familial and residential) characteristics are taken under control. In order to do this we are interested in distinguishing the socio-economic status into three components usually not jointly considered in the literature: human capital (education), financial resources (income), material conditions (housing). How large are socio-economic inequalities in the Italy of the first decade of the new century? Are they mainly attributed to human capital deficits or to financial deficits or jointly to both factors? Do the housing conditions have an autonomous effect? Answering these questions is important to public health intervention strategies: it may in fact lead to more effective interventions to improve health in the whole community, particularly for those who are most vulnerable.

Secondly, we are interested in examining whether the effect of socio-economic factors depends on the individual's residential context. Empirical evidence suggests that health is influenced by local environmental characteristics as well as by personal circumstances (Mitchell et al., 2000; Dominguez-Berjon et al., 2005; Cummings et al., 2005; Basta et al., 2008): individuals living in socio-economically deprived areas are at increasing risk of bad self-rated health or functional impairment, even after controlling for the effects of individual socio-economic status. Thus, environment may have an autonomous impact on the health of individuals: more deprived communities might offer less opportunity for healthy living because of high crime rates, poor housing conditions, environmental pollution, or lack of services. However, what we are interested in here is not the main effect of the context but its interaction with individual socio-economic status. Does the individual's socio-economic status have differentiated effects on health status depending on the territorial context of the respondent? Are people in lower social conditions less healthy if they are living in more disadvantaged environments? Italy is a country with several territorial differences. Those between the North and the South are well known. They are cultural as well as socio-economic. In general, the South shows less favourable conditions than the North with respect to economic, social, and environmental conditions (ISTAT, 2009).

Moreover, in the South the measures of health are worse than in the North, at least at a descriptive level (Ongaro and Salvini, 2009). Our question is, then: do health inequalities due to socio-economic differences operate more in the South than in the North of the country?

In order to answer all these questions, the study makes use of a representative cross-sectional survey carried out by ISTAT in 2004-05. The large sample size allows us to also analyse the data at a sub-national level.

Data and population will be described in the following section. In section 3 we will give a descriptive overview of the magnitude of the socio-economic inequalities in health in Italy. In section 4 we turn to the socio-economic differences, modelling their effects on the health net of other disturbing factors. In the first part we consider only main effects; in the second part we explore whether or not an interaction exists between socio-economic factors and residential context. In the conclusion we present our final comments.

2. Source, data and variables

The different aspects concerning inequalities in the domain of health are analyzed through the Italian health interview survey *Health conditions and recourse to health services* carried out by ISTAT in 2004-2005. Thanks to an agreement among Istat, Regional Authorities and the Department of Health, this wave of the survey can take advantage of a more numerous sample size with respect the usual Italian multipurpose surveys. Using a two-stages stratified sampling design (municipalities and families), the survey has concerned 50,474 families (against the about 24,000 of multipurpose surveys), for 128,040 persons interviewed (ISTAT, 2007b). This large sample size allows reaching the objective of statistical significance of estimations, also at sub-regional level (Caranci *et al.*, 2008).

The survey provides a rich amount of information to investigate several dimensions of population health: from the perception of health condition to the presence of different diseases (acute or chronic), to the presence of disability; from the use of health services to the consumption of drugs and non-conventional cures; from the prevention, physical activities and life styles of individuals to problems relied to the pregnancy.

In this contribution, we will focus on three measures of health of individuals: the perceived health, the multicronicity and the disability.

The first indicator refers to the perception that people have of their health condition, and it allows gathering the concept of health in its comprehensiveness and multidimensionality. The subjective (or perceived) health is measured through

the question suggested by World Health Organisation (WHO) "How is our health in general?" In following analysis we will consider together the categories "very good" , "good" and "fair" on the one side, and "bad" and "very bad" on the other.

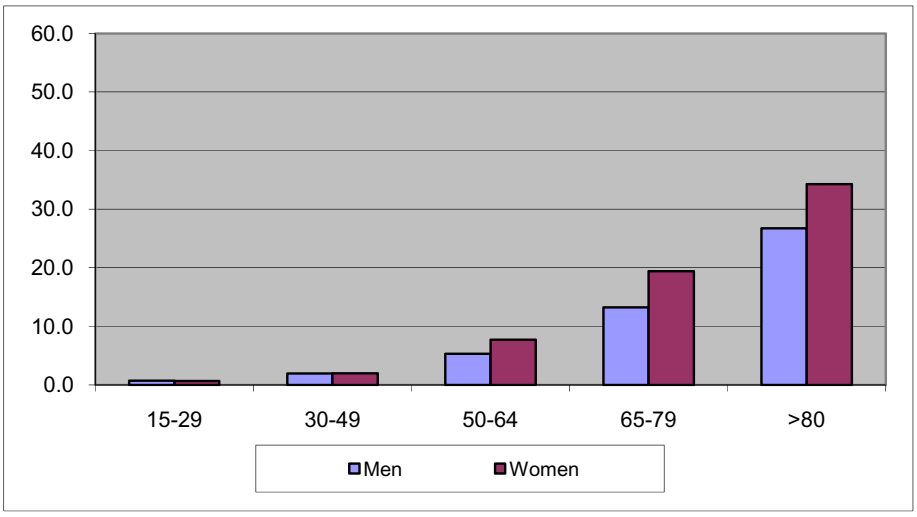
The indicator of multiple chronic diseases identifies individuals suffering from three or more chronic diseases at the same time. Particularly, since the survey collects declarations of individuals about the presence of diseases regardless of a medical certification of the pathology, in a sense also in this case we deal with a subjective indicator of health, even if in a more attenuated way compared to the perception of comprehensive health.

The third measure of health conditions points out the presence of disability. An individual is qualified as disabled if he/she declares to have serious difficulties at least in one of the three main dimensions of health: the physical dimension, the sphere of the activities of daily living, and the domain of communication. Also in this case, the responses, though based on health conditions objectively verifiable, imply a subjective evaluation of the individuals of their own condition.

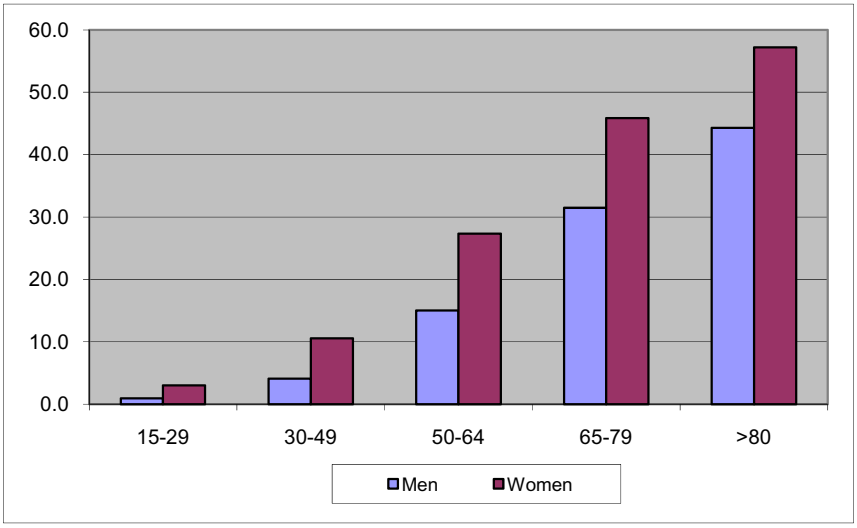
Since the prevalence of bad health conditions according with these measures is very, or relatively, low in youthful and central ages (Figure 1), the analysis we present will focus on adult and elderly people (50 years and over).

Figure 1 (a-b-c) – *Prevalence of poor health (bad o very bad self-rate health; 3 or more chronic diseases; disability) according to sex and age: individuals aged 50 and over.*

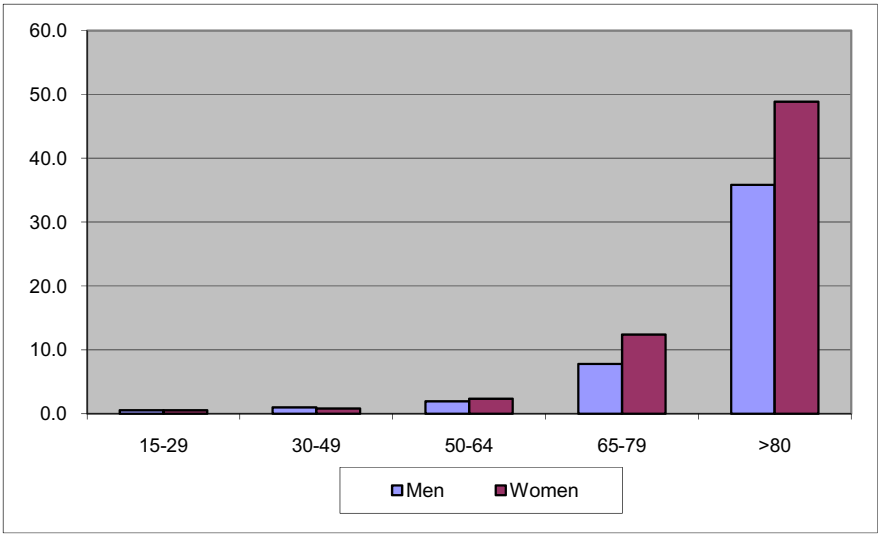
a) Self rated health: percentage of people declaring to be in bad or very bad health



b) Chronic diseases: percentage of people declaring 3 or more chronic diseases



c) Disability: percentage of people with disability



Our purpose is to analyse these three measures of health with respect some socio-economic conditions, both individual and familial, in order to analyse whether and to what extent the Italian population represents a homogeneous picture

or it still presents some inequalities from the point of view of health. The socio-economic conditions of individuals are taken into account using three indicators coming from the survey and referring to the education, the economic situation and the housing condition. The level of education is coded in "low", "medium" and "high" according to the respondent's education is respectively at the primary, low-secondary and high-secondary/university levels. The financial situation is taken into account using the subjective evaluation expressed by individuals about the adequacy of their economic resources. Particularly, the indicator used distinguishes people who judge the total economic resources disposable to their family "very good" or "appropriate" from people who judge them "scarce" or "inadequate". Finally, using a set of six indicators concerning some characteristics of the house[1] , an indicator referring to the housing conditions has been built. Based on preliminary analysis, in this paper we consider the opposition between bad housing conditions (index values from 1 to 6) and very good housing conditions (index value equals to 7).

Since one of the aims of our analysis is to verify the possibility that the socio-economic inequalities interact with the territorial context, we will introduce into the analysis a variable referring to the area of residence of individuals (North, Centre and South).

Other covariates are considered as background factors. The family structure of individuals and their relationships networks represent another contextual factor at *meso*-level that may interact both with socio-economic individual variables and with the *macro* context. There are two variables we deem useful to control in this regard. The first is the marital status, which modalities have been aggregated differentiating between married people and people who do not live in couple (single, separated or divorced, widowed[2]). The second indicator refers to the help potentially available to individuals in case of need. It has been build using the questions concerning the fact that people have (1) relatives, (2) friends or (3) neighbours that they could rely on in case of need. The modalities used in the analysis oppose people who can rely on help of relatives, friends and neighbours on the one side, and people who can rely at most on two of these helps on the other.

---

[1] The index of housing conditions is based on the following characteristics declared by individuals about their house: lack of bathroom, lack of heating system, house too small, presence of humidity stains, house in bad conditions, less than one room per component. In case of presence of all these negative conditions, the index will take value 1, that equals to very bad conditions. The index equals 2 in case of presence of 5 out of 6 negative characteristics, and so on, till a value of 7 whether no one of the negative conditions is present (that equals to very good housing conditions).

[2] The questionnaire construction prevent from identifying people who, though not married, live in couple.

Finally, as usual, the analysis will be controlled for sex and age of individuals.

3. Health status according to individuals' socio-economic conditions

Table 1, which describes the health status of the population aged 50 and over according to the three measures of socio-economic conditions previously presented in section 2, unambiguously shows that there are quite distinct socio-economic differences in health status among Italian people over 50.

The level of education is negatively associated with bad health (tab. 1.a): those with low educational levels feel on average worse than those who have medium education, and those who have medium education feel on average worse than those who have high education. This relationship is observed among both men and women and is substantially independent of the age class of respondents. There are, however, differences according to the type of measure used. The effect of educational level is weak when health is measured by the existence of three or more chronic diseases, but is very evident when it is measured by means of the prevalence of disability or self-rated perception of health. The negative association between educational level and disability is particularly evident at younger ages. Subjects between 50 and 64 years old with low educational level are three times more likely to be disabled with respect to their age-related peers with a high level; in 65-79-year-olds, those with low educational level are twice more likely to be disabled than their peers with high levels. Only in the "oldest old" do relative differences become attentuated, but differences in the prevalence of disability still remain - at least 10-12% between those with high and low educational levels. The strong effect of educational level on health may also be observed when health is measured by the self-rated perception of health. Up to the age of 80, those with high levels have a probability of self-rated poor health which is less than half that of their age-related peers with low levels (among 50-64-year-olds, the percentages of people with poor self-rated health are respectively 3-4% vs. 8-11%; among people aged 65-79 the same percentages are respectively 7-11% and 16-22%). Beyond the age of 80, differences in educational level are attenuated, but still remain considerable (people with self-rated poor health pass from 19-26% for those with the highest educational level to 28-36% with the lowest level. As regards this measure of health, high educational level seems to have a protective effect, so that highly educated subjects report their health as more similar to that of subjects of lower educational level in the next lower age group than to that of subjects of the same age groip with low level of education.

Health differences also emerge between people with differing financial resources (Tables 1.b and 1.c). Independently of subjects' age and gender, those

with low incomes and poor-quality housing have worse health, thus highlighting the risk that they accumulate problems of various kinds. Those with insufficient finances or poor housing also have higher probabilities of having 3+ chronic diseases, of being disabled, or of self-rating poor health. Among individuals aged 50-64, those with low incomes have double the probability of being disabled and treble the probability of poor health, compared with those who declare sufficient incomes. In 65-79-year-olds, those with low incomes have double the probability of being disabled and double that of poor health, compared with those who have sufficient means. Over the age of 80, health differences by income are attentuated, but do not disappear (all health measures show differences of prevalence of illness or disability of about 10-14%).

Briefly, men and women of low educational level, with insufficient incomes and poor housing state that they more often have several chronic diseases, but mainly that they are disabled and in poor health. At great ages, the phenomenon tends to recede, presumably because of the prevalence of a fragility effect due to age, which operates independently of socio-economic status. It is in fact little probable that this result reflects bias arising from exclusion of the institutional population: although in the fourth age (subjects over 80) the percentages of people in care increase significantly (6-7%) with respect to younger people (1%) [1], the proportions are still too low to explain the fall in health differentials by socio-economic status in the highest age group.

The associations found between state of health on one hand and educational level, income and housing on the other, are, however, certainly spurious. Low educational levels also often imply low incomes and poor housing. In addition, it is possible that persons who are fragile from the viewpoint of personal resources also experience other situations of family and environmental difficulties which may lie at the origin of their poor state of health. In the following section, we estimate the effects of each socio-economic factor on health regardless of other confounding variables.

Table 1 (a-b-c) – *Health status of women and men by age, and socio-economic characteristics (level of education, economic resources, housing conditions). Italy 2004-05.*

a) Level of education

Educational level	50-64		65-79		80+	
	Men	Women	Men	Women	Men	Women
	individuals with bad or very bad self rated health (%)					
Low	7,8	10,7	16,1	21,5	27,7	35,7
Medium	5,2	6,7	10,6	17,1	29,9	30,6
High	3,4	4,3	7	10,9	19,1	25,8
	individuals with 3+ chronic diseases (%)					
Low	18,5	31	33,3	46,8	45,1	57,4
Medium	13,8	25,8	29,5	45,2	46,8	59,2
High	13,2	23,4	27,7	41,2	38,3	51,6
	individuals with disability (%)					
Low	3,4	3,5	9,3	13,8	36,8	51
Medium	1,7	1,9	6,3	10,7	41,4	42,3
High	1	1,1	4,7	6,9	25,9	38,2

b) Economic resources

Economic resources	50-64		65-79		80+	
	Men	Women	Men	Women	Men	Women
	individuals with bad or very bad self rated health (%)					
very good or appropriate	3,4	5,2	9,6	14,2	21,0	29,1
scarce or inadequate	10,5	13,5	20,6	28,1	36,6	42,1
	individuals with 3+ chronic diseases (%)					
very good or appropriate	12,6	24,4	27,9	41,5	40,7	53,1
scarce or inadequate	21,7	34,0	38,5	53,1	50,5	63,2
	individuals with disability (%)					
very good or appropriate	1,4	1,7	6,2	10,0	31,6	44,5
scarce or inadequate	3,5	3,8	11,1	16,3	43,2	55,3

c) Housing conditions

Housing conditions	50-64		65-79		80+	
	Men	Women	Men	Women	Men	Women
	individuals with bad or very bad self rated health (%)					
bad	9,0	11,4	18,6	28,5	33,8	40,5
good	4,1	6,6	11,7	16,9	24,7	32,4
	individuals with 3+ chronic diseases (%)					
bad	18,0	30,7	34,7	52,4	47,6	62,2
good	14,0	26,3	30,6	44,1	43,3	55,6
	individuals with disability (%)					
bad	3,4	3,6	10,5	18,8	44,3	58,4
good	1,5	2,0	7,1	10,6	33,5	45,9

4. Modelling the effects of socio-economic factors

4.1 *Main effects models*

A synthesis of the relationship is achieved through the construction and assessment of logistic regression models where the dependent variable is, alternatively, the self-perceived health, the multi chronic morbidity and the disability. For every previously defined health indicator, we have verified the effect of some socio-economic covariates, such as level of education, financial situation (using the subjective evaluation expressed by individuals about the adequacy of their economic resources) and the index of housing conditions.

Moreover, we have included in the models two variables as proxies of aid-individual network. The first is marital status, to take into account the presence of a partner, and the second is an indicator that refers to the help potentially available to individuals in case of need. It has been build using the questions concerning the fact that people have relatives, friends or neighbours that could help respondents in case of need.

Finally place of residence (geographic macro-region) has been included in the model. The aim of this choice is two-fold: to suggest some considerations about the

role of cultural context and to take into account eventual differences of sanitary services. The analyses have been controlled for sex and age group.

In the first model, where the perceived health status is the dependent variable, the entire set of background variables is significant (table 2). It is interesting to underline that people who judge their financial condition "scarce" or "inadequate" present the highest risk of bad health perception, second only to the age effect. Level of education appears negatively correlated with a bad health perceived condition, confirming the gradient high-low level of education. Also housing conditions seem to influence health perception in the expected way: the better housing conditions are, the lower is the risk of declaring to be in a bad health.

Also when the dependent variable of the logistic regression model is represented by multiple chronic conditions (table 3), all the background variables are significant (even if the effect of marital status is quite weak, it remains significant at 10% significance level). The effect of the intermediate level of education is not significantly different from the high one, while the direct association between risk of three or more chronic diseases and perception of scarce financial resources persists. The effect of the housing conditions remains, but it seems very weak. In this case, nevertheless, coefficients are lower with respect to health perception, evidencing less marked relations of the indicator of multiple chronic diseases with considered socio-economic factors.

Regarding the results of the application of the logistic regression model to disability (table 4), the effect of level of education is significant both for intermediate and lower levels, with respect to the higher one. Equally, it seems significant the role of perception of inappropriate economic condition and the indicator of housing situation, although the coefficients are lower than in the case of self-perceived health. In synthesis, the whole set of results of the main effects models describe a quite consistent picture. Socio-economic variables present a direct effect on health indicators, regardless the indicator used, highlighting that people belonging to the lower social classes are more prone to manifest bad health conditions.

Net of demographic characteristics (gender and age group) and family and neighbouring situation (marital status, presence of relatives or friends), relationships are stronger and more evident than for objective ones (multiple chronic diseases and disability).

Geographic context seems to influence in the same way all the aspects of the health status. Taking Northern regions as reference category, those who reside in the Centre-South declare less favourable health conditions, in particular people living in the South. While relations regarding multichronicity are less evident, subjective perception underlines a dichotomous situation in geographical differences of health conditions in Italy.

Table 2 – *Results of logistic regression model: poor self rated health.*

Covariates	*Coefficients*	*O.R.*
Gender		
Men (reference)	0,00	1,00
Women	0,31	1,38
Age Group		
50-64 (reference)	0,00	1,00
65-79	0,86	2,35
>80	1,62	5,06
Marital Status		
Married (reference)	0,00	1,00
Single/Separated/Divorced/Widow	0,12	1,13
Aid-network (relatives and friends)		
Not all (reference)	0,00	1,00
All	-0,32	0,73
Residence		
North (reference)	0,00	1,00
Centre	0,40	1,49
South	0,48	1,62
Level of education		
High (reference)	0,00	1,00
Medium	0,32	1,37
Low	0,59	1,81
Economic situation		
Adequate (reference)	0,00	1,00
Scarse	0,73	2,07
Housing condition		
Very good (reference)	0,00	1,00
Not very good	0,27	1,31

Note: All coefficients are significant at 1% significance level; italic character indicate coefficients significant at 10% significance level

Table 3 – *Results of logistic regression model: multiple chronic diseases.*

Covariates	*Coefficients*	*O.R.*
Gender		
Men (reference)	0,00	1,00
Women	0,61	1,83
Age Group		
50-64 (reference)	0,00	1,00
65-79	0,79	2,21
>80	1,24	3,46
Marital Status		
Married (reference)	0,00	1,00
Single/Separated/Divorced/Widow	*0,04*	*1,04*
Aid-network (relatives and friends)		
Not all (reference)	0,00	1,00
All	-0,14	0,87
Residence		
North (reference)	0,00	1,00
Centre	0,23	1,25
South	0,16	1,17
Level of education		
High (reference)	0,00	1,00
Medium	*0,04*	*1,04*
Low	0,19	1,20
Economic situation		
Adequate (reference)	0,00	1,00
Scarse	0,44	1,55
Housing condition		
Very good (reference)	0,00	1,00
Not very good	0,12	1,13

Note: All coefficients are significant at 1% significance level; italic character indicate coefficients significant at 10% significance level

Table 4 – *Results of logistic regression model: disability.*

Covariates	*Coefficients*	*O.R.*
Gender		
Men (reference)	0,00	1,00
Women	0,27	1,31
Age Group		
50-64 (reference)	0,00	1,00
65-79	1,44	4,24
>80	3,23	25,27
Marital Status		
Married (reference)	0,00	1,00
Single/Separated/Divorced/Widow	0,48	1,62
Aid-network (relatives and friends)		
Not all (reference)	0,00	1,00
All	-0,14	0,87
Residence		
North (reference)	0,00	1,00
Centre	0,15	1,17
South	0,48	1,62
Level of education		
High (reference)	0,00	1,00
Medium	0,32	1,37
Low	0,59	1,81
Economic situation		
Adequate (reference)	0,00	1,00
Scarse	0,40	1,50
Housing condition		
Very good (reference)	0,00	1,00
Not very good	0,25	1,29

Note: All coefficients are significant at 1% significance level; italic character indicate coefficients significant at 10% significance level.

4.2 *Are the effects socio-economic factors depending on the residential context?*

Starting from previous results, we tested in the logistic regression models also the interaction effects between the geographical region of residence (North, Centre, South) and some socio-economic factors. Among the latter, perception of economic resources, used as proxy of income, provides some interesting results.

Table 5 shows the main effects together with the interaction ones. Whereas the interaction is not significant for the Centre, the coefficient for the South confirms the importance of the impact of the context-income relationship in determining the perception of a bad health status. The coefficient is even higher for multiple chronic diseases and disability.

The differences resulting from the estimated models may be highlighted also comparing the different "profiles" in terms of odd-ratio (Figure 2 a-b-c). With "North-good economic resources" as category of reference, the gradient North-South seems evident: in the southern regions, people declare a worse self-rated health than in the northern ones, with central regions in the middle of the ranking. Controlling for region, people belonging to the lower income groups present a worse health status, regardless of the indicator used.

Table 5 – *Impact of geographical region and economic condition on health indicators. Results of logistic regression models: main and interaction effects.*

Covariates	Poor self rated health		Multiple chronic diseases		Disability	
	Coeff.	Signif.	Coeff.	Signif.	Coeff.	Signif.
Scarce economic resources (refer.: adequate)	0,811	0,000	0,554	0,000	0,479	0,000
Centre (ref.: North)	0,429	0,000	0,262	0,000	0,151	0,013
South (ref.: North)	0,554	0,000	0,243	0,000	0,555	0,000
Centre*scarce econ. resources	*-0,069*	*0,367*	*-0,107*	*0,062*	*-0,002*	*0,984*
South* scarce econ. resources	-0,160	0,011	-0,237	0,000	-0,174	0,019

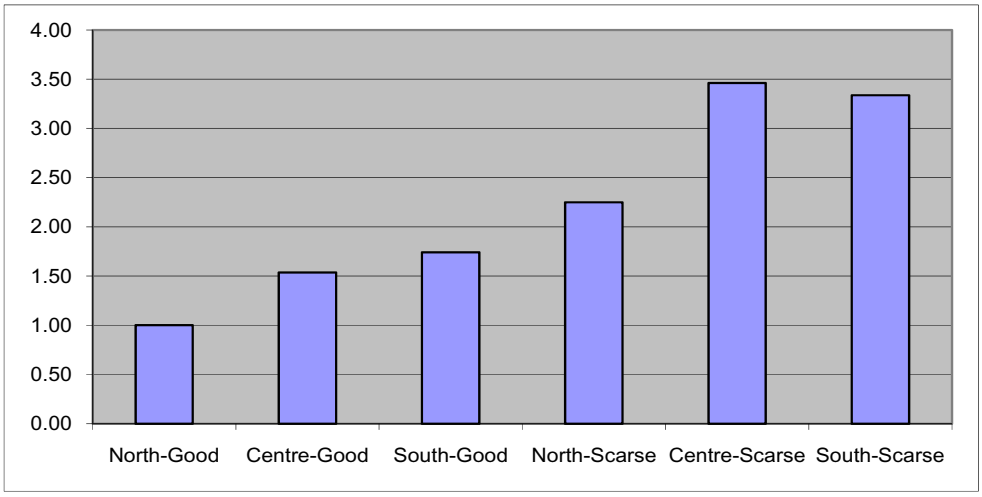
Note: All coefficients except to those in Italic character are statistically significant at least at the level of 0.05. Results are controlled for the main effects of sex, age group, marital status, aid-network, level of education and housing condition.

The significance of the interaction between region of residence and economic resources proves that the residential context has different effects both on health status and on the perception of one's own economic condition. This confirms the hypothesis that it exists a "contextual effect" which, in the case under study, seems to impacts negatively on people living in the Centre and in the South of Italy.

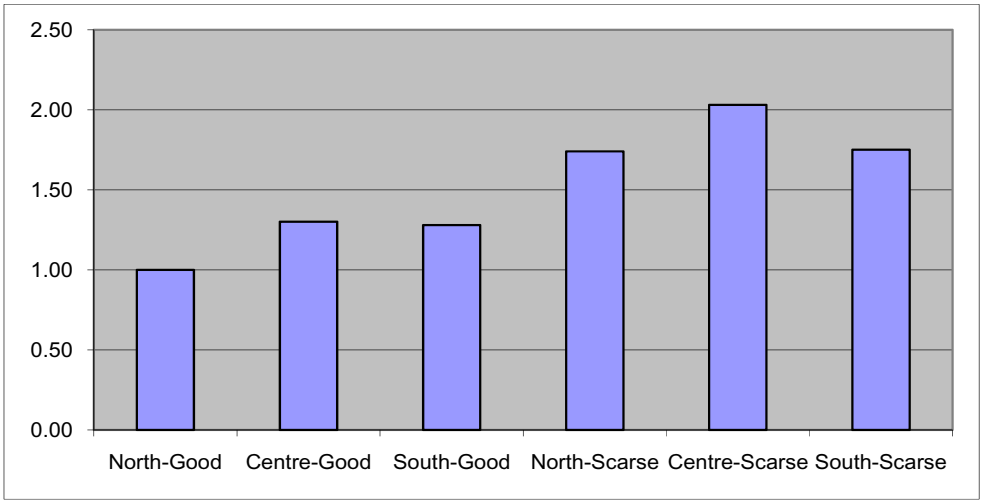
These results offer a synthetic way to compare the impacts of the differences existing among territorial units with regards many aspects. In fact, geographic region synthesizes many factors, first of all the availability of sanitary services. But we can assume that also subjective aspects, such as psychological and cultural characteristics of people living in the different regions are into play, even if they are not ever easy to detect and to interpret.

Figure 2 (a-b-c) – *Impact of geographical region and economic condition on health indicators. Results of logistic regression models for the interaction between geographical region and perception of economic resources: odds-ratios.*

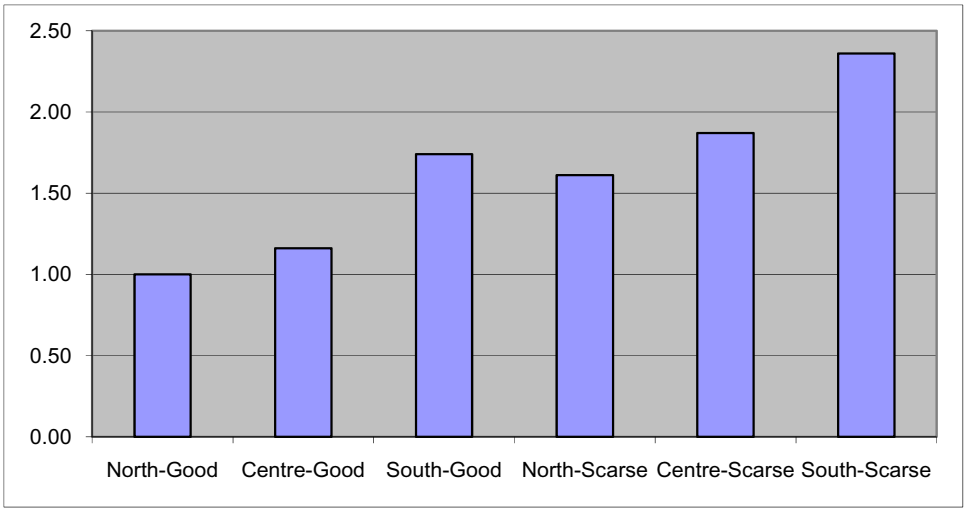
(a) Self rated health



(b) Multiple chronic diseases



(c) Disability



Note: The model includes interactions between perception of economic resources (good, scarce) and geographical region of residence (North, Centre and South); reference category are North and Good economic resources. Results are controlled for the main effects of sex, age group, marital status, aid-network, level of education and housing condition.

5. Concluding remarks

The first result of this study is that, still in the first decade of the new century, Italy presents health differences depending on individual socio-economic status. Once controlled for personal characteristics, there is an autonomous effect of socio-economic status, which worsens the health status of less privileged people. This is the case especially for the self-rated health but, in a lower measure, it happens also when considering disability and multi chronic morbidity. This means that, on the one hand, there is still a lot of scope for improvements in health of Italian population, and, on the other hand, that the public health system – as observed in other developed countries – is not able to answer in a differentiated way to the specific and various health needs of the population. For example, it is not able to offer proper services and make successful interventions to control, at least, part of the behavioural risk factors (smoking, dietary habitus which are different from one group to the other) to which most of the socio-economic differences of health seems to be due.

The second result is that each component representing the individual's socio-economic condition (educational level, financial conditions, housing assets) has an autonomous impact on the individual's health status. Socioeconomic factors have, indeed, different effects on the various health measures: financial conditions present a higher impact on both self-rated health and multiple chronic diseases, whereas educational level has a large impact on disability. More important, they operate on health with cumulated effects. Scarce health conditions presented by the most vulnerable segments of population not only depend on low income levels, but also on poor human capital. Even housing conditions have an autonomous impact on health status: consequently, interventions aiming to enhance both the hygiene and the state of repair of buildings may contribute to the improvement of health conditions. How and through which mechanisms these different factors are operating, and why they have different effects on the different dimensions of health, may be a subject of future research.

A further result of the study is that environmental context influences autonomously the health of individuals. As documented by empirical literature, also in Italy individuals living in socio-economically deprived areas are at increasing risk of perceiving a poor health status. This is reflected into a North-South gradient, such that people living in the North of the country declare lower rates of morbidity and disability than people living in the South.

Moreover, the study documented that the context influences the way through which the individual's socio-economic status (specifically financial resources) affects his/her health status. Surprisingly, in contrast with our hypotheses, people living in less developed regions of the South experience lower health inequalities

due to income differences than those living in the more developed Northern regions. This issue is scarcely studied in literature and needs further analyses. However, this result seems to be consistent with other studies that, using macro approaches, suggest the presence of a negative association between the level of health and health inequalities (Bommier, Stecklov, 2002). Considering that i) similar results were found for Italy in a recent study using regions as unit of analysis (Mazzuco, 2009) and that ii) the worse health status is mainly observed in the southern regions, we may hypothesize that health inequalities in the South with respect to the North might be explained, at least in part, by the different effects that individual income has on health in the two geographical areas. If this evidence were confirmed by other studies, it supports the call to make specific attention to the less advantaged socio-economic groups, when general health conditions are improving.

Finally, the study suggests further more technical investigations. The analysis was carried out using subjective indicators of both health and economic status. Self-rated health has indeed been proven to be an excellent indicator: for example, it is a reliable predictor of survival; moreover, self-assessment of level of functional disability works well since individuals associate it with independence, autonomy and ability. However, we do not forget that both indicators are also an outcome of public attitudes and cultural wisdom. Moreover, little is known about the robustness of measures of economic conditions based of self-assessment. Further investigations on the validity and reliability of these indicators may be useful also in order to better interpret the differentiated (main and interaction effects) effects of the geographical context on the individual's health status. Other analyses should also be addressed to better isolate the influence of the context. The study used the macro regions as proxy of the context, considering that in Italy they are different with respect to socio-economic, environmental, cultural, and health care aspects (see for example the competences on health cure that the law delegates to the Regions). However, we do not know which are the specific factors that come into play when the geographical context seems to have a significant impact on the health of population. Deeper analysis using *area level variables* (Costa et al., 2003) are requested to better understand the mechanism through which the context influences the health net of other individual factors.

References

Basta N.E., Matthews F.E., Chatfield M.D., Brayne C., MRC-CFAS (2008), Community-level socio-economic satus and cognitive and functional impairment in the older population, The European Journal of Public Health, 18(1): 48-54.
Bommier A., Stecklov G. (2002), Defining health inequality: why rawls succeeds where social welfare theory fails, Journal of Health Economics, 21 (3), 497-513.
Caranci et *al*. (2008), *Metodi per l'analisi dei dati*, in Quaderni di Monitor - Approfondimenti sull'indagine multiscopo Istat Salute 2005, Agenzia nazionale per i servizi sanitari regionali, Anno VII, n. 22.
Commission on Social Determinants of Health (2007), A Component Framework for Action on the Social Determinants of Health, Draft, April 2007, www.who.int/social_determinants/resources/csdh_framework_action_05_07.pdf.
Costa G., Marinacci C., Caiazzo A., Spadea T. (2003), Individual and contextual determinants of inequalities in health: the Italian case, International Journal of Health Services,, 33, 635-667.
Costa-Font J. (2008), Housing assets and the socio-economic determinants of health and disability in old age, Health Place, 14 (3): 478-91.
Cummings S., Stafford M., Macintyre S., Marmot M., Ellaway A. (2005), Neighbourhood environment and its association with self-rated health: evidence from Scotland and England, Journal of Epidemiology and Community Health, 59: 207-13.
Dominguez-Berjon F., Borrell C., Rodriguez-Sanz M., Pastor V. (2005), The usefulness of area-based socio-economic measures to monitor social inequalities in health in Southern Europe, Eur. Jou. of Public Health, 16, 1: 54-61.
Giskes K.,Turrel G., Patterson G.,Newman B. (2002), Socio-economic differencesin fruit and vegetable consuption among Australian adolescents and adults, Public Health Nutrition, 5: 663-69.
ISTAT (2007a),
http://www.istat.it/salastampa/comunicati/non_calendario/20070302_00/testointegrale.pdf.
ISTAT (2007b), Nota metodologica dell'Indagine "Condizioni di salute, fattori di rischio e ricorso ai servizi sanitari" – Anno 2005, Statistiche in breve, Roma.
ISTAT, 2009, Rapporto Annuale. La Situazione del Paese nel 2008, ISTAT, Roma.
Lanzieri G. (2008), Population and social conditions, Statistics in focus, 81/2008, Eurostat.
Mackenbach J.P. (1992), Socio-economic health differences in the Netherlands: a review of recent empirical findings, Social Science and Medicine, 34(3): 213-26.
Mitchell R., Gleave S., Bertley M., Wiggins D., Joshi H. (2000), Do attitude and area influence health? A multilevel approach to health inequalities, Health Place, 6: 67-79.
Mazzuco S. (2009), Condizioni di salute nelle regioni italiane, in Ongaro F., Salvini S. (eds), op. cit.
Ongaro F., Salvini S. (eds), 2009, Rapporto sulla Popolazione, Salute e sopravvivenza - Gruppo di Coordinamento per la demografia, Il Mulino, Bologna.
Van der Heiden J.H.A., Demarest S., Tafforeau J., Van Oven H. (2003), Socio-economic differences in the utilisation of health services in Belgium, Healh Policy, 65(2):153-65.

Volkers A.C., Westert G.P., Schellevis F.G. (2007), Health disparities by occupation, modified by education: a cross-sectional population study, BMC Public Health, Aug. 8; 7; 196.

SUMMARY

Referring to the Italian population, a general improvement is observed in the field of health in recent years, even if elements of heterogeneity still exist. Besides the well-known inequalities of age and gender, persistent differences in health conditions depending on socio-economic status are observed, at least at a descriptive level. This study intends to analyse more in depth the relationship between health status and socio-economic conditions in light of the most recent data on health conditions in the Italian population. In particular, our aim is twofold. Firstly, we intend to verify to what extent the differences observed at a descriptive level persist after other individual or contextual (familial and residential) characteristics are taken under control. Secondly, we are interested in examining whether the effect of socio-economic factors depends on the individual's residential context. Our question is, then: do health inequalities due to socio-economic differences operate more in the South than in the North of the country?
The analysis of the socio-economic inequalities in health status presented in this contribution, is carried out through logistic regression models focusing on three measures of individual health – the perceived health, the multicronicity and the disability – computed making use of the representative cross-sectional survey "*Health conditions and recourse to health services*", carried out by ISTAT in 2004-05. The large sample size and the sample design of this survey allow performing the analysis also at a sub-national level.

Fausta ONGARO, Università di Padova
Silvana SALVINI, Università di Firenze

Rivista Italiana di Economia Demografia e Statistica Volume LXIII nn. 1-2 – Gennaio-Giugno 2009

ESTIMATION OF POVERTY INDICATORS IN ITALY: A SMALL AREA M-QUANTILE BASED APPROACH

Monica Pratesi, Caterina Giusti, Nicola Salvati

1. Introduction

The estimation of local poverty measures, such as the European Laeken indicators, is nowadays an important objective to address since policy makers should dispose of information referred to appropriate domains.

In any area of interest, the knowledge of the cumulative distribution function of an income variable represents an important source of information for the living conditions in that area. From the cumulative distribution function of the household disposable income many quantities (e.g. the median income, the income quantiles) and monetary poverty indicators (e.g. the Head Count Ratio or at-risk-of-poverty-rate) can be computed. Moreover, it is important to note that the knowledge of the cumulative distribution function of the income allows to also estimate the proportion of population whose income is immediately under or above a given poverty threshold. Thus, the estimation of traditional monetary poverty indicators accompanied by the estimation of the cumulative distribution function of the income variables of interest can fulfil a more detailed picture of poverty.

However, the estimation of all these quantities at a detailed geographical level is complicated by the available survey information. In fact, data coming from the EU-SILC survey, the most complete and valuable source of information to produce poverty and living condition estimates in Italy, can be used to produce accurate estimates only at the NUTS 2 level (that is, regional level). Thus, to satisfy the increasing demand from official and private institutions of statistical estimates on poverty and living conditions referring to smaller domains (LAU 1 and LAU 2 levels, Local Administrative Units 1 and 2, that is Provinces and Municipalities), there is the need to resort to small area methodologies.

Small area estimation techniques are employed when sample data are insufficient to produce accurate direct estimates in the domains of interest. The idea of these methods is to use statistical models to link the survey variable of interest with covariate information that is also known for units not in the sample. Among these models traditional linear mixed models (Rao, 2003) are the most popular. More recently Chambers and Tzavidis (2006) developed a M-estimator for

small area parameters based on the quantiles of the distribution of the study variable. There are many features of this novel approach, which makes it appealing and promising in small area estimation. Firstly the method avoids assumption of normality on the response variable; in addition, due to the presence of an influence function in M-estimator, it can control the influence of outlying observations. Moreover it is straightforward extendible to take into account also geographical information on small areas and it can be adapted to situations when the relationship between the quantiles of the response variable and the covariates is not linear. Many are the application fields where these features are important. In our case, estimation of poverty indicators and poverty mapping can surely take advantage of these. In fact income has not a normal distribution and the presence of outlying observations often make it difficult to adapt model on it. In addition the distribution of income often shows variations due to local factors often well captured by inserting geography in the model. Finally it is not far from reality that the effect on income of some traditional explaining covariates (imagine, for instance the age of the head of the household) be not linear.

These features will not be described in depth here. References to these properties are in several recent works as the paper by Pratesi et al. (2009), Salvati et al. (2009) and Giusti et al. 2009).

Here the focus is on the use of a bias adjusted estimator of the cumulative distribution function based on an M-quantile model. Our objective is the estimation of the cumulative distribution function of the household equivalised income, and thus of some income quantiles in each small area. Particularly we focus on the estimation of some poverty measures, such as the Head Count Ratio, for the Provinces of three Italian Regions, namely Lombardia, Toscana and Campania. The aim is to analyze the potential poverty dissimilarities present inside each Region, better investigating at the same time also the so-called "North-South divide" characterizing the Italian territory. For this purpose we combine data coming from the EU-SILC survey 2007 with those from the Population Census 2001.

2. Theory

Let $\mathbf{x}_i$ be a known vector of p auxiliary variables for each population unit j in small area i and assume that information for the variable of interest y (e.g., the household income) is available only for the sampled units. The aim is to use these data to estimate various quantities in each small area.

A popular approach for this purpose is to use mixed effects models with random area effects to model household income. Given the so-called unit level nested error

regression model (Battese *et al.*, 1988), the Empirical Best Linear Unbiased Predictor (EBLUP) of the mean of *y* in small area *i* is:

$$\hat{m}_i^{EBLUP} = N_i^{-1}\left[\sum\nolimits_{j\in s_i} y_j + \sum\nolimits_{j\in r_i} \hat{y}_j\right] \tag{1}$$

where $\hat{y}_j = \mathbf{x}_j^T\hat{\boldsymbol{\beta}} + z_j\hat{u}_i$ are the values predicted under the assumed model; s_i denotes the n_i sampled units in area *i*: r_i denotes the remaining $N_i - n_i$ units in the area ; $\hat{\boldsymbol{\beta}}$ and $\hat{u}_i$ are obtained by substituting an optimal estimate of the covariance matrix of the random effects into respectively the best linear unbiased estimator of $\boldsymbol{\beta}$ and the best linear unbiased predictor of u_i. It is important to associate to estimator (1) a measure of its variability. The Mean Squared Error (MSE) of (1) and its estimate are obtained following the results of Kackar and Harville (1984) and Prasad and Rao (1990). Details and formulas can be found in Rao (2003, Chapter 7).

Recently, Chambers and Tzavidis (2006) have developed another approach to small area estimation based on the quantiles of the conditional distribution of the study variable given the covariates (Breckling and Chambers, 1988). The q^{th} M-quantile $Q_q(x;\psi)$ of the conditional distribution of *y* given *x* satisfies:

$$Q_q(\mathbf{x}_{ij};\psi) = \mathbf{x}_{ij}^T\beta_\psi(q) \tag{2}$$

where ψ denotes the influence function associated with the M-quantile. For specified *q* and continuous ψ, an estimate $\hat{\beta}_\psi(q)$ of $\beta_\psi(q)$ is obtained via an iterative weighted least squares algorithm. When (2) holds, the bias adjusted M-quantile predictor of m_i is:

$$\hat{m}_i^{MQ/CD} = N_i^{-1}\left[\sum_{j\in s_i} y_j + \sum_{j\in r_i}\mathbf{x}_j^T\hat{\beta}_\psi\left(\hat{\theta}_i\right) + \frac{N_i - n_i}{n_i}\sum_{j\in s_i}\left(y_j - \hat{y}_j\right)\right] \tag{3}$$

where $\hat{y}_j = \mathbf{x}_j^T\hat{\beta}_\psi(\hat{\theta}_i)$ is a linear combination of the auxiliary variables and $\hat{\theta}_i$ is an estimate of the average value of the M-quantile coefficients of the units in area i (Tzavidis and Chambers, 2007). The MSE of the estimator (3) can be estimated analytically as suggested in Chambers *et al.* (2007).

While there are many alternative estimators of the small area mean, the estimators of the distribution function have not yet been developed at small area

level. A useful starting point is the so-called Chambers and Dunstan biased adjusted estimator of the small area distribution function in the presence of outliers (Tzavidis et al., 2008a). This is defined as:

$$\hat{F}_i^{CD}(t) = N_i^{-1}\left[\sum_{j\in s_i} I(y_j \le t) + n_i^{-1}\sum_{j\in s_i}\sum_{k\in r_j} I\left\{\hat{y}_k + (y_j - \hat{y}_j) \le t\right\}\right]. \quad (4)$$

The p^{th} quantile m_{pi} of the distribution of y in area i can be estimated by the solution to:

$$\int_{-\infty}^{m_{pi}} d\hat{F}_i^{CD}(t) = p. \quad (5)$$

We can note that by substituting properly $\hat{y}_j$ in (4) one can define M-quantile or mixed model versions of the CD-based distribution function estimator as well as corresponding estimators of the within area quantiles of y. Also nonparametric versions of the CD-based distribution function estimator and geographically based estimators of it can be found defining properly $\hat{y}_j$ by means of non parametric or semiparametric M-quantile regression and geographically weighted regression. More details on these extensions can be found in Pratesi et al. (2009), Giusti et al. (2009) and Salvati et al. (2008).

The estimation of the MSE of (4) for the case of linear M-quantile models is under study and the first results can be found in (Tzavidis et al., 2009).

3. An application to the estimation of poverty in three Italian Regions

The European Survey on Income and Living Conditions (EU-SILC) is conducted yearly in every Member State of the European Union to produce comparable and timely estimates on poverty and living conditions of individuals and household, both in a cross-sectional and longitudinal perspective.

Regions are planned domains of the Italian EU-SILC, thus estimates referring to this geographical level (corresponding to Nomenclature of Territorial Units for Statistics NUTS-2) are published every year. The regional samples are based on a stratified two stage sample design: Municipalities are the Primary Sampling Units (PSUs), while the households are the Secondary Sampling Units (SSUs). Provinces

and Municipalities (LAU1 and LAU2 levels) are instead unplanned domains: some Provinces may have very few sampled Municipalities, and many Municipalities are not even included in the sample at all. Direct estimates at these geographical levels may therefore have large errors or they may not even be computable. Thus, to produce estimates of poverty and living conditions referring to Italian Provinces and Municipalities there is the need to resort to small area estimation techniques.

In each selected household each member older than 15 years is interviewed with an individual questionnaire, and one member of the household, usually the head of the household, is also interviewed with a household questionnaire. These questionnaires together fulfil a lot of information on the living conditions of the selected households, mainly in terms of micro-data on income, housing and living condition, social exclusion.

The target of our case study is the estimation of the mean household income, of the Head Count Ratio (HCR) or incidence of poverty and of some income quantiles for the Provinces of three Italian Regions: Lombardia (Northern Italy), Toscana (Central Italy) and Campania (Southern Italy). Data on the household equivalised income, on some household characteristics and on individual characteristics of the head of the household in the three Regions are available from the EU-SILC survey 2007. The same covariate information is available from the Census 2001 for all the households living in Lombardia, Toscana and Campania. The aim is not only to evaluate the distribution of the income inside the three Regions, but also to get a picture of the poverty and living conditions inequalities characterizing the Italian territory.

In this application we use an M-quantile CD estimator (3) with the following covariates: the marital status of the head of the family (six levels), the working position of the head of the household (four levels), the education of the head of the household (ten levels), the gender of the head of the household (male/female) and the mean house surface at area level (in square meters). The small areas are the 12 in Lombardia (11 Provinces plus the Municipality of Milano), 11 in Toscana (10 Provinces plus the Municipality of Firenze) and six in Campania (5 Provinces plus the Municipality of Napoli), for a total of 29 small areas. Log-transformation of household income has not yet been considered at this stage of the work, to avoid the possible bias and the complications of the back-transformation on the MSE estimation of the small area estimators (Chambers and Dorfman, 2003).

The estimates of the Head Count Ratio (HCR) and of its lower and upper confidence levels, estimated using a bootstrap technique, and the estimates of the mean household income and of the corresponding standard errors are in Table 1. The mean income estimates referring to Lombardia, Toscana and Campania are also represented in Figures 1, 2 and 3 respectively. In the figures the darker color of an area corresponds to a worst situation of poverty, that is a lower mean income.

Finally, Table 2 reports the estimates of some income quantiles in the areas of interest.

The first evident results of our analyses is the higher incidence of poverty in the areas of Campania, a Region in Southern Italy: for this Region the estimates of the HCR, the percentage of households below the poverty line (9504 Euros, corresponding to the 60\% of the median income), are in the range 26-44%, while for Lombardia (Northern Italy) and Toscana (Central Italy) the ranges of the HCR are 5-19% and 9-27% respectively. Also the estimated values for the mean income suggest a gap between these three Italian Regions (Table 1).

Nevertheless, for both the estimates of interest we can notice a certain variability inside the three Regions. For example (Table 1), several areas in Lombardia have an estimated HCR between 12% and 19%, included the value of the Municipality of Milano. In Campania we can notice a gap in the income estimates referring to the Northern areas (Caserta, Benevento and Napoli Provinces), with lower income estimates with respect to the Southern areas of the Region; however the same gap characterize only the HCR estimate referring to the Province of Benevento. For the Toscana Region we can notice that the higher incidence of poverty as well as the lower mean income estimate refer to the Province of Massa-Carrara, in the North of the Region. For this area the value of the HCR is comparable to the lower HCR values we observe in the Campania Region.

Table 1 – *Small areas estimates using the M-quantile CD model: population and small area sizes, Head Count Ratio, HCR lower and upper limits.*

Areas	N_i	n_i	HCR	HCR lower limit	HCR upper limit	Mean income	Mean s.e.
Massa-Carrara / Toscana	80810	96	0,26	0,21	0,32	14842,84	676,21
Lucca / Toscana	146117	131	0,21	0,16	0,26	16690,43	799,59
Pistoia / Toscana	104466	129	0,14	0,11	0,18	19255,50	1134,13
Firenze Province / Toscana	216531	326	0,11	0,09	0,13	18644,50	454,15
Livorno / Toscana	133729	115	0,15	0,11	0,20	18738,69	940,46
Pisa / Toscana	150259	136	0,15	0,10	0,19	19167,14	878,97
Arezzo / Toscana	123880	145	0,13	0,10	0,16	19414,74	1056,74
Siena / Toscana	101399	116	0,12	0,09	0,16	20928,79	1078,15
Grosseto / Toscana	87720	59	0,17	0,12	0,23	17874,16	1138,76
Prato / Toscana	83617	117	0,12	0,08	0,16	18097,09	691,09
Firenze Municip. / Toscana	159724	119	0,12	0,09	0,16	22203,33	1320,28
Varese / Lombardia	320899	253	0,10	0,08	0,13	21928,52	1297,09
Como / Lombardia	205963	153	0,16	0,13	0,22	19361,44	1130,66
Sondrio / Lombardia	69817	41	0,19	0,12	0,28	16894,17	1625,11
Milano Prov. / Lombardia	957305	543	0,12	0,10	0,14	20265,20	510,59
Bergamo / Lombardia	375778	219	0,17	0,13	0,20	19212,59	829,83
Brescia / Lombardia	437706	216	0,18	0,14	0,21	16921,99	572,86
Pavia / Lombardia	211786	60	0,15	0,10	0,23	22053,49	4053,48
Cremona / Lombardia	135321	75	0,16	0,10	0,21	17222,00	882,97
Mantova/ Lombardia	146249	234	0,13	0,10	0,16	18546,50	656,01
Lecco / Lombardia	121321	103	0,11	0,07	0,16	20281,30	1127,10
Lodi / Lombardia	77978	62	0,11	0,07	0,17	17986,44	995,64
Milano Mun. / Lombardia	588197	255	0,12	0,09	0,14	23876,47	1113,60
Caserta / Campania	279684	155	0,38	0,33	0,43	12056,16	611,14
Benevento / Campania	102441	70	0,43	0,35	0,52	12109,72	1003,61
Napoli Prov. / Campania	631523	596	0,40	0,38	0,43	12104,09	351,01
Avellino / Campania	152340	84	0,36	0,29	0,43	13609,70	990,96
Salerno/ Campania	359080	191	0,32	0,28	0,37	13534,64	522,29
Napoli Mun. / Campania	337787	221	0,26	0,22	0,30	16399,39	626,67

Figure 1 – *Estimates of the mean equivalised household income - M-quantile CD Estimator, Lombardia Region.*

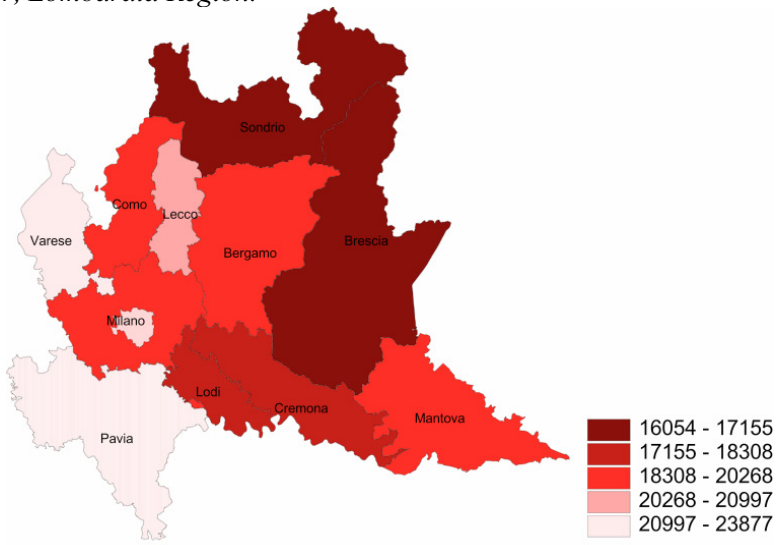


Figure 2 – *Estimates of the mean equivalised household income - M-quantile CD Estimator, Toscana Region.*

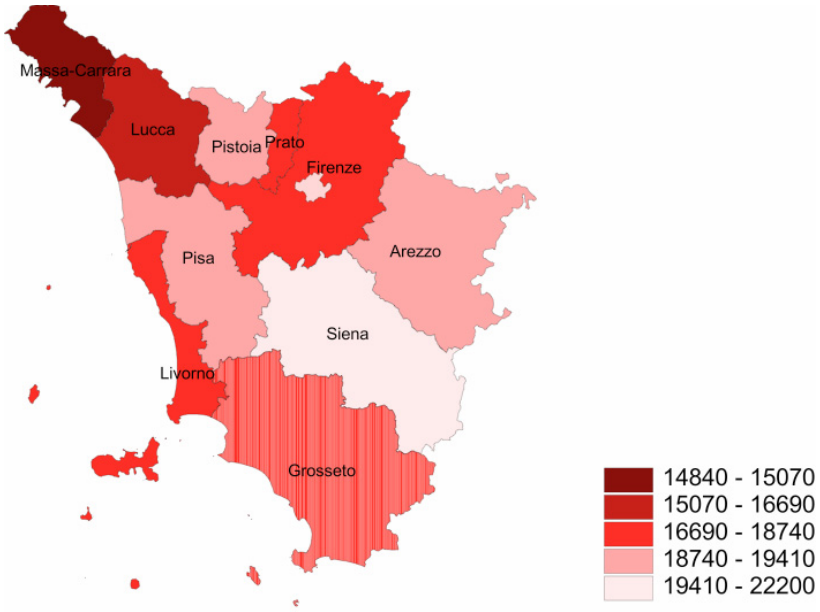
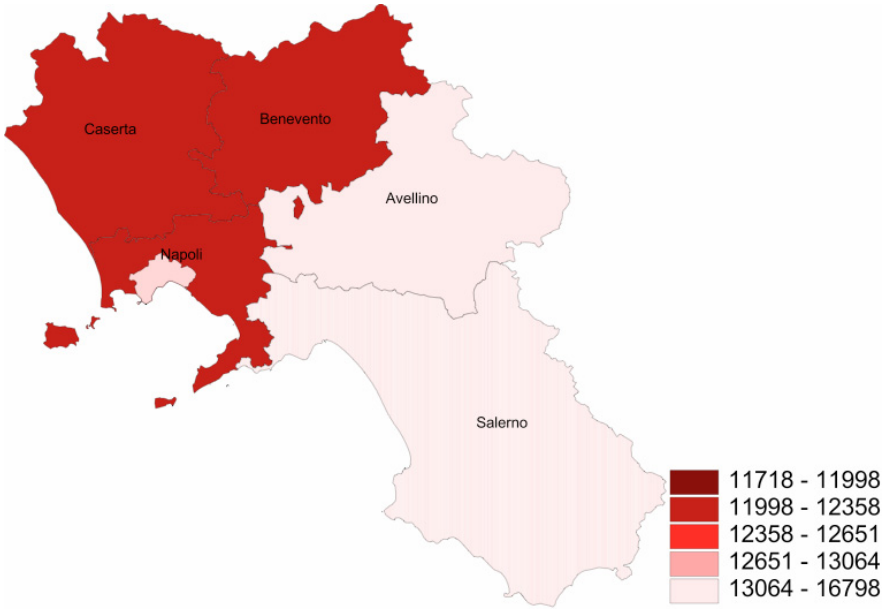


Figure 3 – *Estimates of the mean equivalised household income - M-quantile CD Estimator, Campania Region.*



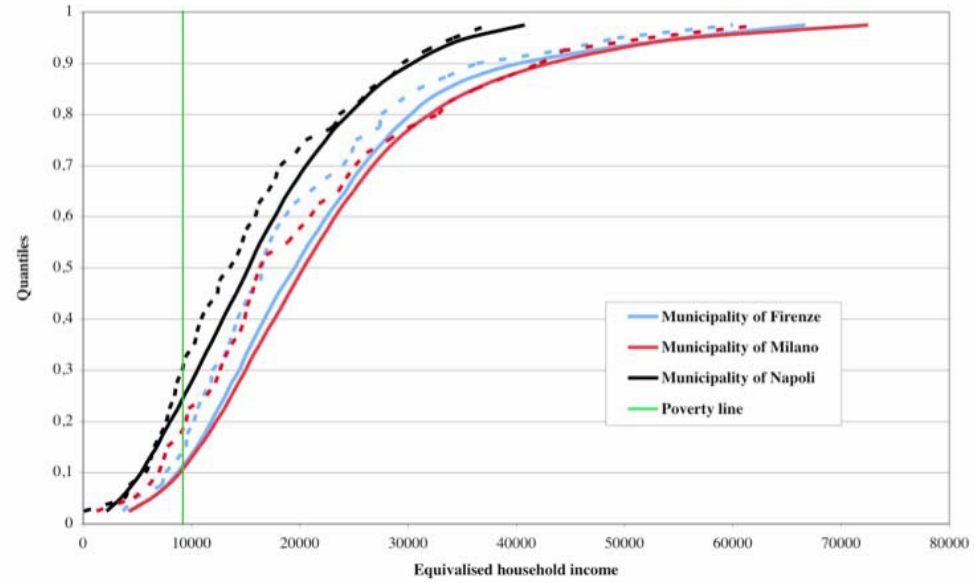
The Municipality of Milano, though characterized by a relative high HCR in the Lombardia Region, is in the class of higher estimated mean income, and it also has the median and the upper quantile estimates in Lombardia (see Figure 1, Tables 1 and 2). This suggests that for income values over the poverty line the cumulative distribution function of this area is above all the other estimated cumulative distribution functions in Lombardia. A similar behaviour characterized the other two big Municipalities, Firenze in Toscana and Napoli in Campania, though in these cases the estimated HCRs are always in the lower class of the corresponding Region. If we compare the cumulative distribution functions estimated for the three Municipalities (Figure 4 and Table 2) we can appreciate the gap between the Municipality of Napoli and the other two areas, and we can see that the quantiles of the Municipality of Milano are slightly higher than those of the Municipality of Firenze for income values above the poverty line. The direct estimates computed through the Horvitz-Thompson estimator (dashed lines in Figure 4) are not always consistent with the model based ones, especially in the centre of the income distributions. These results suggest the relevance of the information that can be

obtained through the estimation of the cumulative distribution function of the household income in the areas of interest.

Table 2 – *Small areas estimates using the M-quantile CD model: quantiles of the household equivalised income.*

Areas	*q=0.1*	*q=0.25*	*q=0.50*	*q=0.75*	*q=0.90*
Varese / Lombardia	9386,46	13148,25	17646,15	24105,93	33694,71
Como / Lombardia	7823,42	11592,31	16501,40	22992,27	33327,84
Sondrio / Lombardia	7466,26	10656,90	15113,02	20461,93	26047,79
Milano Prov./ Lombardia	8731,20	13192,02	18269,21	24382,96	33250,05
Bergamo / Lombardia	7407,39	11490,67	16682,27	23879,87	35312,04
Brescia / Lombardia	7463,28	10973,61	15567,05	21559,37	28221,11
Pavia / Lombardia	7727,44	11662,38	16815,47	24158,36	30846,34
Cremona / Lombardia	8014,57	11612,81	15944,49	21565,67	29013,89
Mantova/ Lombardia	8703,20	12167,10	16950,97	22972,99	29916,88
Lecco / Lombardia	8926,71	13201,13	18592,04	25331,02	34951,54
Lodi / Lombardia	8851,35	12909,56	17385,32	22686,72	28789,72
Milano Mun. / Lombardia	8921,27	13658,16	20336,33	28928,33	43009,34
Massa-Carrara / Toscana	5927,56	9302,67	13897,29	19656,76	25352,02
Lucca / Toscana	6698,79	10341,86	15518,95	21849,49	28174,46
Pistoia / Toscana	8100,64	11771,77	16466,57	22676,13	35314,37
Firenze Prov. / Toscana	9081,97	12815,30	17265,50	23021,86	30129,86
Livorno / Toscana	7637,14	11932,00	17169,57	24035,98	31825,37
Pisa / Toscana	7715,23	12310,66	17651,00	25020,65	31989,20
Arezzo / Toscana	8526,24	12627,70	17321,35	22980,46	31214,38
Siena / Toscana	8616,74	13122,44	19234,35	26341,03	34390,34
Grosseto / Toscana	7369,55	11342,94	16950,41	24261,85	30086,36
Prato / Toscana	8735,63	12888,64	17351,54	22823,14	28960,56
Firenze Mun. / Toscana	8685,82	13104,39	19497,05	27816,45	40169,88
Caserta / Campania	3955,83	6966,30	11865,18	17680,72	24114,55
Benevento / Campania	4047,04	7013,11	10658,07	15280,13	21469,55
Napoli Prov. / Campania	3764,81	6709,65	11190,69	16569,87	22664,62
Avellino / Campania	4163,12	7469,51	12163,71	18480,82	24591,62
Salerno/ Campania	4650,65	8145,19	13018,57	18407,38	24263,83
Napoli Mun. / Campania	5360,07	9301,79	15256,73	22367,41	30363,27

Figure 4 – *Estimated cumulative distribution functions of household equivalised income - M-quantile CD Estimator, Municipalities of Milano, Firenze and Napoli.*



4. Concluding remarks

The results presented in this work suggest that the availability of data from the EU-SILC survey and from the Population Census, together with the use of small area methodologies, can fulfil information on poverty in areas where reduced survey samples prevent from resorting to direct survey estimators.

In particular, the estimation of the cumulative distribution function of the household income allows not only to compute monetary poverty indicators such as the Head Count Ratio, the median or the income quantiles in each small area, but also to "track" the distribution of the income. This is particularly important since the same incidence of poverty (e.g. the same HCR) could correspond to different income distributions, that is to a different proportion of population immediately below the poverty line, and thus at risk to become poor.

In the case-study we focused on data referring to three Italian Regions, Lombardia, Toscana and Campania. The estimated poverty indicators confirm the existence of a gap between the level of the mean and median household income in the North/Centre and in the South of Italy. However, other interesting results emerged form the analyses. For example, the Municipality of Milano is characterized by high estimated upper quantiles of the income, but the incidence of

poverty is nevertheless a relevant one. A similar result, though less evident, characterizes the Municipalities of Firenze and Napoli.

The estimates referring to the Provinces are only a first step towards a more in-depth study of the income and poverty disparities in the three Regions. In next steps of our work we will compute the estimates at a more detailed geographical level, such as the Municipalities. These results could also give new insides in the so-called "North-South divide" of Italy, as suggested by the results at provincial level.

Next developments of our analyses will also regard the estimation of non-monetary indicators of poverty (Cheli and Lemmi, 1995), and the feasibility to consider the log transformation of income, in order to enhance the fitting of the income model and to exploit the significance of the covariates. At the same time, more work has to be done on the back-transformation to apply appropriate formulas for the mean squared errors.

Further developments will include also the specification of nonparametic and geographically based models for the relations of the quantiles of the income distribution with suitable covariates.

Finally, as the estimator of the income cumulative distribution function is a very useful tool to follow the behavior of the distribution of income at the small area level, we will concentrate on the estimation of its mean squared error, in order to track a confidence interval around the cumulative distribution function line.

Acknowledgments

This work has been supported by the project SAMPLE "Small Area Methods for Poverty and Living Condition Estimates" financed by the European Commission under the 7thFP. For more information please visit the project web-site: www.sample-project.eu

References

Battese, G., Harter, R. and Fuller, W. (1988). An Error-Components Model for Prediction of County Crop Areas using Survey and Satellite Data. *Journal of the American Statistical Association*, 83, 28-36.

Bradolini A., Saraceno C. (2007), Introduzione, in *Povertà e Benessere. Una geografia delle disuguaglianze in Italia*, a cura di Bradolini A. and Saraceno C. Il Mulino.

Breckling J. and Chambers R. (1988). M-quantiles. *Biometrika*, 75, 761-71.

Chambers R., Tzavidis N. (2006) M-quantile models for small area estimation, *Biometrika*, 93, 255-268.

Chambers, R., Chandra, H. and Tzavidis, N. (2007). On Robust Mean Squared Error Estimation for Linear Predictors for Domains. CCSR Working paper 2007-10. Cathie Marsh Centre for Census and Survey Research, University of Manchester.
Chambers R., Dorfman A.H. (2003). Transformed Variables in Survey Sampling. S3RI Methodology Working Papers, M03/21, Southampton Statistical Sciences Research Institute, University of Southampton, UK.
Cheli, B. and Lemmi, A. (1995). A Totally Fuzzy and Relative Approach to the Multidimensional Analysis of Poverty. *Economic Notes*, 24, 115-134.
Eliers, P. and Marx, B. (1996). Flexible Smoothing using B-splines and Penalized Likelihood (with comments and rejoinder). *Statistical Science*, 11, 1200-1224.
European Commission (2006). *Description of SILC Database Variables: Cross-sectional and Longitudinal*. Version 2004.1 from 25-05-06. European Commission – Eurostat.
Foster J., Greer J., Thorbecke E. (1984). A class of decomposable poverty measures. *Econometrica*, 52, 761-766.
Giusti, C., Pratesi, M. and Salvati, S. Small area methods in the estimation of poverty indicators: the case of Tuscany, to apper in *Politica Economica*.
Giusti, C., Pratesi, M., and Salvati, S. (2009). Estimation of poverty indicators: a comparison of small area methods at LAU1-2 level in Tuscany. *Abstract Book NTTS - Conferences on New Techniques and Technologies for Statistics*, Brussels, 18-20 February.
Kackar R.N. and Harville D.A. (1984), Approximations for standard errors of estimators for fixed and random effects in mixed models. *Journal of the American Statistical Association*, *79*, 853-862.
Opsomer J.D., Claeskens G., Ranalli M.G., Kauermann G., Breidt F. J. (2008) Nonparametric small area estimation using penalized spline regression, *Journal of the Royal Statistical Society: Series B*, 70, 265-286.
Prasad, N.G.N. and Rao, J.N.K. (1990). The Estimation of the Mean Squared Error of Small Area Estimators. *Journal of the American Statistical Association*, 85, 163-171.
Pratesi M., Ranalli M.G., Salvati N. (2008) Semiparametric M-quantile regression for estimating the proportion of acidic lakes in 8-digit HUCs of the Northeastern US, *Environmetrics*, 19, 687-701.
Pratesi M., Ranalli, M.G. and Salvati, N. (2009) Nonparametric M-quantile regression using penalized splines. *Journal of Nonparametric Statistics*, 21, 287-304.
Rao, J.N.K. (2003). *Small Area Estimation*. New York: Wiley.
Ruppert, D., Wand, M.P. and Carroll, R. (2003). *Semiparametric Regression*. Cambridge University Press, Cambridge, New York.
Salvati, N., Tzavidis, N., Pratesi, M. and Chambers, R. (2008) Small Area Estimation Via M-quantile Geographically Weighted Regression. [paper submitted for publication, available upon request]
Salvati, N., Ranalli, M.G. and Pratesi, M. (2009). Small area estimation of the mean using nonparametric M-quantile regression: a comparison when a linear mixed model does not hold. To appear in *Journal of Statistical Computation & Simulation*.
Särndal C.E., Swensson B., Wretman J.H. (1992). *Model Assisted Survey Sampling*. New York, Springer-Verlag.

Tzavidis N., Salvati N., Pratesi M., Chambers R. (2008). M-quantile Models with Application to Poverty Mapping. *Statistical Methods & Applications*, 17, 393-411.
Tzavidis N., Marchetti S., Chambers R. (2009). Robust prediction of small area means and quantiles. To appear in *Australian and New Zealand Journal of Statistics*.

SUMMARY

Estimation of poverty indicators in Italy: a small area M-Quantile based approach

The estimation of local poverty measures, such as the European Laeken indicators, is nowadays an important objective to address since policy makers should dispose of information referred to appropriate domains. In this paper we resort to M-quantile small area estimation methods to estimate the Head Count Ratio, the mean income and some income quantiles in the Provinces of three Italian Regions (LAU1 level), namely Lombardia, Toscana and Campania. Moreover, to fulfil a more detailed picture of poverty, we estimate the cumulative distribution function of the income in the three principal Municipalities of the Regions, Milano, Firenze and Napoli.

Monica PRATESI, Associate Professor in Statistics, Department of Statistics and Mathematics Applied to Economics, University of Pisa. Contact e-mail: m.pratesi@ec.unipi.it.
Caterina GIUSTI, Research Fellow in Statistics, Department of Statistics and Mathematics Applied to Economics, University of Pisa. Contact e-mail: caterina.giusti@ec.unipi.it.
Nicola SALVATI, Researcher in Statistics, Department of Statistics and Mathematics Applied to Economics, University of Pisa. Contact e-mail: salvati@ec.unipi.it.

Rivista Italiana di Economia Demografia e Statistica Volume LXIII nn. 1-2 – Gennaio-Giugno 2009

THE EVOLUTION AND DISTRIBUTION OF INCOME IN A LIFE-CYCLE PERSPECTIVE

Giambattista Salinari, Gustavo De Santis

1. Geometric random walks

How does income evolve over time? Gibrat (1931) offers a solution with his model, which, as we show below, accurately mimics the dynamic of several phenomena, including income (Sutton 1997; Gabaix 1999; Davis and Shorrocks 2000; Mitzenmacher, 2003), and works better than its critics originally thought (cf. Kalecki 1945; Champernowne 1953; Rutherford 1955; Hart 1976; Shorrocks 1976; Armatte 1995; Neal and Rosen 2000).

Gibrat's original formulation focuses on the evolution of labor earnings, described as a pure stochastic process: the logarithm of income in year t+1 is given by the logarithm of income in year t plus a random shock. The evolution of log-income can therefore be describe as a pure random walk, and the stationary distribution of incomes that eventually emerges is the log-normal distribution.

Now, two features of Gibrat's model appear surprising, and did in fact attract criticism: 1) labor incomes evolve at random, and; 2) this (random) evolution is not stationary, because its mean may vary and its variance increases constantly. The former characteristic becomes more easily acceptable if one considers that, although the effect on income of a few independent variables is relatively predictable (e.g. coming from a wealthy family, marrying the "right" person, getting a degree, etc.), the possible causes of change are virtually innumerable, and so are their interactions. The interplays of these forces is so complex that it evades our understanding, and appears as "chance" to our eyes: chance shapes the distribution of incomes, produces inequalities and, on the left tail of the distribution, generates poverty.

Gibrat's model describes a dynamic process (the *evolution* of income), but the lack of longitudinal data has frequently forced researches to focus on its static consequences (the *distribution* of income at a given time). And the fact that the cross-sectional distribution of incomes (supposedly log-normal) did not fully conform to a log-normal curve is perhaps has often been taken as indirect evidence that the model does not work well, in practice.

What this article sets out to do is to use Gibrat's model in a dynamic sense, applying it to the *evolution* of individual income, measured longitudinally, with panel data. By adopting a life-cycle perspective (Jenkins 2000) we get back to the original spirit of Gibrat's model: the log-income of an individual of age *a+1* is given by the log-income of the same individual at age *a* plus a random shock. We show that the model works well within cohorts of dependent workers, and also at any given time, because the cross sectional distribution of incomes at time t need not conform to a log-normal distribution: rather, to a mixture of log-normal distributions, one for each cohort of workers, with varying seniorities.

As mentioned, Gibrat's model implies that the variance of incomes (i.e., inequality) must increase over time. However, empirical data have normally displayed either the opposite trend, especially in the past (Kalecky 1945) or, more recently, no trend at all (Atkinson 1997; Gottshalk 1997; Gotshalk and Smeeding 2000). How do we reconcile this empirical finding with the theoretical expectation? Once again, this apparent contradiction depends on the difference of perspective: Gibrat's model forecasts an increase in the variance of incomes with age (i.e within cohorts), but not necessarily over time: as older cohorts, with high variance in incomes, get out of the observation (because they die, or leave the labor market), they are replaced by younger ones, with lower variance. This renewal of generations halts, and may even revert, the tendency towards an increase in income variance over time.

Gibrat's model has thus far not been extensively used with panels: this type of data was not available when the model was "fashionable", and later on (i.e. during the past 20 years or so) the model has been relatively neglected. But now that panels are becoming more easily available, it may be worthwhile to look more closely at how this model performs in practice, in a life-cycle perspective. We intend to show basically two things:

a) that Gibrat's model works generally well, and, at closer scrutiny, passes most of the critics it received;
b) that there remain a few relatively minor deficiencies - not noticed before, to the best of our knowledge -, which may justify the introduction of a slightly amended version of the model. This is the "modified Gibrat's model" that we present in section 4.

In order to compare the theoretical consequences predicted by the model with the actual evolution of individual income, we use the *ECHP*, or European Community Household Panel[1]. Note that, for reasons of homogeneity and consistency with the assumptions of the model, we restrict our observations to the

[1] For more information, please consult the ECHP webpage at http://circa.europa.eu/irc/dsis/echpanel/info/data/information.html.

labor earnings of Italian males, who are dependent workers for the entire period under consideration. This restriction leaves us with slightly more than 1,000 cases. The period covered by the ECHP spans over the years 1993 to 2000, but we decided to retain only the years 1995-2000, in order to avoid the distortions produced by the economic crisis of 1993[2].

2. Gibrat's model: longitudinally and cross sectionally

Gibrat's model works as follows: let a stand for "age since the beginning of work" (or seniority), t for time (calendar year) and Y for the distribution of labor earnings, so that $Y_{a,t}$ indicates the distribution of labor earnings in year t of those whose seniority is a. Besides, let $R_{a,t}$ (=1+$r_{a,t}$) stand for a random variable that describes the evolution, from t to t+1, of the labor income of the employed with a years of seniority, so that

$$Y_{a,t+1} = Y_{a,t}\, R_{a,t} \tag{1}$$

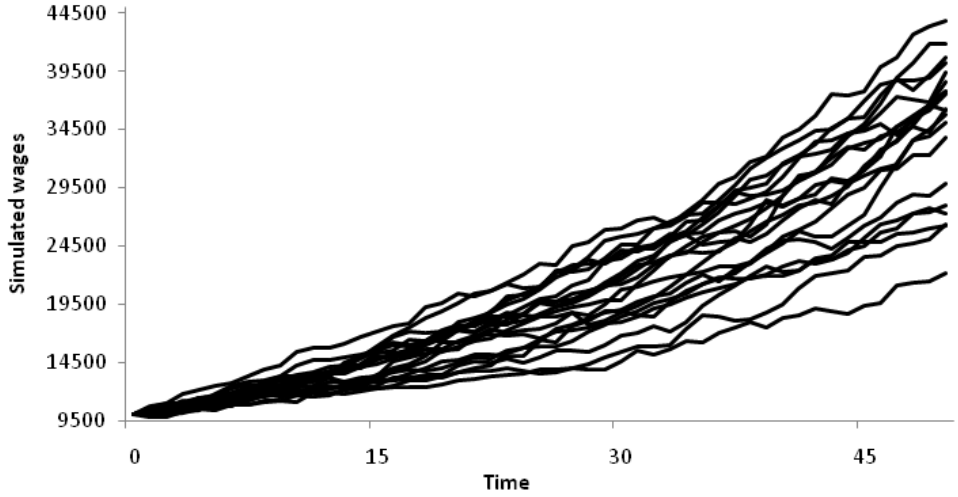
Assuming homogeneity of the population (i.e. no subgroups) and over time - so that R_0, R_1, ..., R_k = iid(μ,σ^2) -, a cohort of workers whose labor incomes were originally distributed as Y_0, will see an evolution of their income distribution of the following type:

$$Y_a = Y_0 \times \prod_{i=0}^{a-1} R_i \tag{2}$$

Figure 1 shows an example of what this means in practice: 20 income trajectories, starting exactly from the same point (10 thousand Euros, in this example) and evolving according to eq. (2), end up relatively far from each other at the end of the 50-year simulation period: between 20 and 45 thousand Euros.

[2] Results do not change qualitatively if the whole period is considered, but they tend to lose some of their statistical significance.

Figure 1 – *Simulation of a Gibrat-type evolution for 20 individual incomes.*



Note: 20 incomes, originally equalling 10,000 Euros, evolve according to model (1)-(2), where ln(R)~N(1.025; 0.022^2).
Source: authors' simulation.

The logarithmic transformation of (2) is

$$\ln Y_a = \ln Y_0 + \sum_{i=0}^{a-1} \ln(R_i) \tag{3}$$

where $\sum_{i=0}^{a-1} \ln R_i = \ln\left(\frac{Y_a}{Y_0}\right) \sim \mathrm{N}$ tends to a normal distribution as a increases, because of the theorem of the central limit.[3]

Cross sectionally, the density function of income in year t is given by the weighted sum of the income density function of individuals with seniority 0, 1,…, a:

$$f(Y_t) = w_{0,t} f(Y_{0,t}) + w_{1,t} f(Y_{1,t}) + \ldots + w_{a,t} f(Y_{a,t}) \tag{4}$$

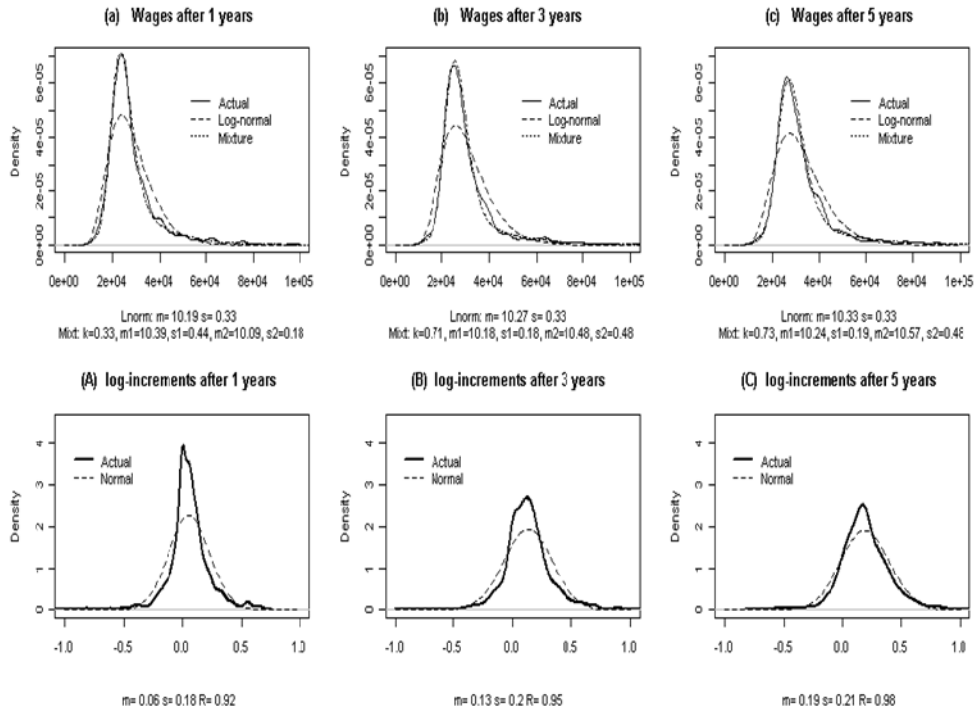
where $f(Y_{a,t})$ indicates the density function of income of individuals having seniority a in year t, and $w_{a,t}$ indicates the proportion of individuals of seniority a

[3] Note, however, that $\ln(Y_a)$ need not be normally distributed, because this also depends on the initial distribution Y_0 (Kalecki 1945).

with respect to the total working population. As a increases, f($Y_{a,t}$) tends towards log-normality, so that f(Y_t) results in a mixture of (quasi) log-normal density functions. Equation (4) links the income distribution in a given epoch with the age structure of the working population. If we assume time homogeneity (the labor earnings of the several cohorts in the population evolve according to the same model) then the income distribution $Y_{a,t}$ of the older workers will show a greater variance than that of the younger ones. This implies, incidentally, that when the working population ages, the variance of the income distribution should increase.

Figure 2 describes the evolution of wages of the Italian dependent workers from 1995 to 2000: panels *a, b*, and *c* show the distributions of wages as of 1996, 1998, and 2000, both the actual and mixture model one. The fit is (obviously) better, virtually perfect, with a mixture model merging two lognormal distributions than with just one lognormal model distribution.

Figure 2 – *Evolution of labor income (A, B, C) and labor income distributions (a, b, c). Italy, 1995 to 2000: empirical data and Gibrat's model.*



Note: optimization criterion: maximum likelihood
Source: Own elaborations on ECHP data.

Below, panels *A*, *B*, and *C* of Figure 2 show the log-increments (log Y_t/Y_0) of incomes with respect to the base year, 1995, for *t*=1996, t=1998 and t=2000, that is 1, 3 and 5 years later, respectively. Gibrat's model (eq. 3) suggests that the log-increments should get closer and closer to a normal distribution as t increases, which conforms to reality.

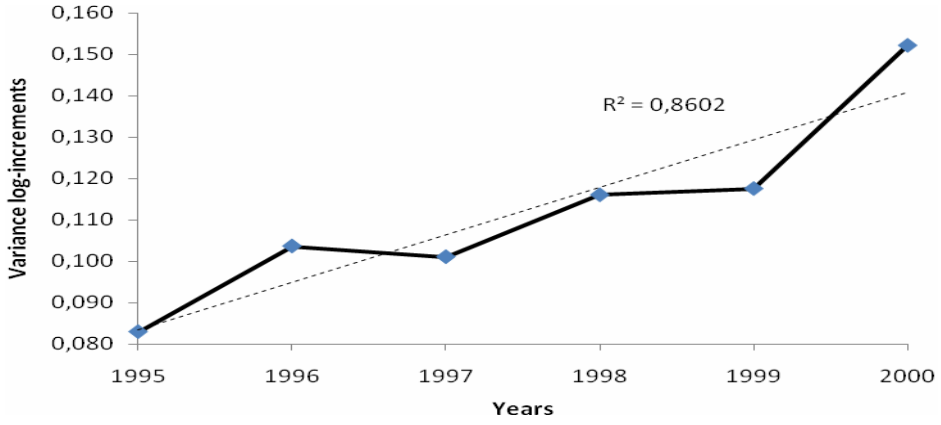
Let us now focus on the variable $\ln(Y_a/Y_0)$, i.e., log increments of income. Both its expected value and its variance should vary linearly with *a* (the time interval considered), because, from eq. (3) we obtain:

$$E\left[\sum_{i=0}^{a-1} \ln R_i\right] = \sum_{i=0}^{a-1} E\left[\ln R_i\right] = a \times \mu_R \tag{5}$$

$$V\left[\sum_{i=0}^{a-1} \ln R_i\right] = \sum_{i=0}^{a-1} V\left[\ln R_i\right] = a \times \sigma_R^2 \tag{6}$$

where μ_R and σ_R^2 are the average and the variance of $\ln(R_i)$. Figure 3 shows that this is actually the case.

Figure 3 – *Variance of the log-increments of income (Italy, dependent workers; 1995-2000).*



Source: Own elaborations on ECHP data

Notice that, in Gibrat's model, the average and the variance of the relative increments Y_a/Y_0 (without logarithmic transformation) depend on each other. Let μ_Y and σ_Y be the average and the standard deviation of the (log-normal) variable Y_a/Y_0; let μ_L and σ_L be the average and the standard deviation of the (normal)

variable $\ln(Y_a/Y_0)$; and, finally, let con μ_R and σ_R be the average and the standard deviation of $\ln(R_i)$. It can be proven (see, e.g., Mitzenmacher 2003) that:

$$\mu_Y = e^{\mu_L + \sigma_L^2/2} \tag{7}$$

$$\sigma_Y^2 = (e^{\sigma_L^2} - 1)e^{2\mu_L + \sigma_L^2} \tag{8}$$

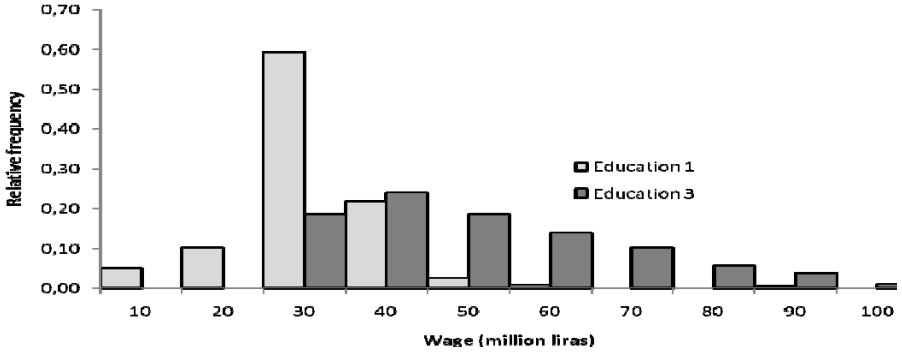
Using (5) and (6), these equations can be transformed into

$$\mu_Y = e^{a(\mu_R + \sigma_R^2/2)} \tag{9}$$

$$\sigma_Y^2 = (e^{a\sigma_R^2} - 1)e^{a(2\mu_R + \sigma_R^2)} \tag{10}$$

In short, the larger the variance of the increments $\left(\sigma_R^2\right)$, the larger the average (μ_Y), and vice-versa. This can explain why those subgroups whose income increases faster than average typically have also a comparatively large variance (see, e.g., Figure 4).

Figure 4 – *Wage distribution among workers with low (1) and high (3) education.*



Note: Workers with low education (education 1) are 572; those with high education (education 3) are 108. During the time interval 1995-2000, the highly educated experienced a mean growth of incomes as high as R_3=1.387, and their standard deviation in 2000 was 24 million liras. During the same period, the mean growth of the incomes of the scarcely educated workers was R_1=1.340, and their standard deviation in 2000 was 9 million liras (estimates refer to nominal wages).
Source: Own elaborations on ECHP data.

Gibrat's model throws some light on the process of income (and, therefore, also social) mobility. In order to see this, let us imagine that we want to "follow" the income evolution of a cohort of persons. When they start to work, their seniority is 0, by definition, but we first observe them only later, at seniority a and we can follow them up to seniority $a+k$. The appendix shows that the covariance between the income distributions at seniorities a and $a+k$ is:

$$\text{cov}\,[\ln(Y_a);\ln(Y_{a+k})] = a\sigma^2 \tag{11}$$

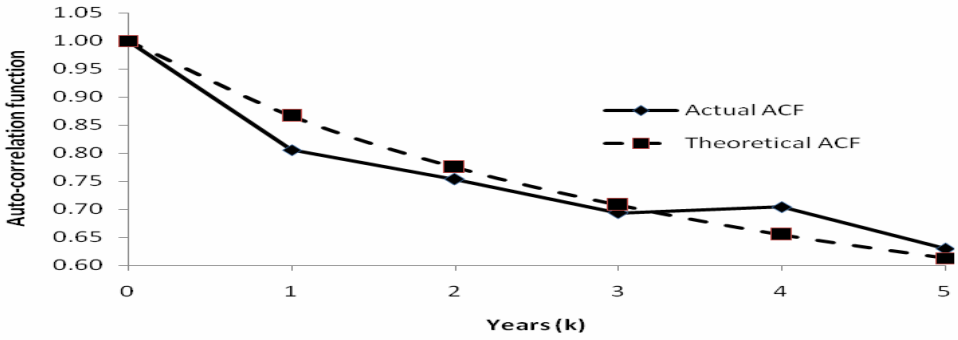
The autocorrelation function is therefore:

$$cor\left[\ln Y_a, \ln Y_{a+k}\right] = \sqrt{\frac{a}{a+k}} \tag{12}$$

which is always positive, less than one, and tends to 0 (highest economic mobility) as k increases (see also Figure 5); but tends to 1 (no economic mobility) as a increases (cf. appendix 2).

In short, as the members of a cohort get older (greater a), income increases (eq. 5) and so do variance and inequality (eq. 6), while economic mobility slows down (eq. 11): the wage structure gets more and more rigid. Moreover, every shock in income dynamics (a poverty spell, for example) produces a permanent effect on the income evolution (because the process is non-stationary).

Figure 5 – *Autocorrelation function of dependent labour incomes (Italy, 1995-2000).*



Note: The empirical auto-correlation function (ACF) has been estimated for the period 1995-2000 by calculating the correlation coefficients between the wages in 1995 (year 0) and in the subsequent years. The theoretical autocorrelation function are the values predicted by eq. (12). The estimate of the parameters of eq. (12) has been obtained through a non-linear regression of the values of the empirical auto-correlation. The R^2 coefficient between the empirical and the theoretical autocorrelation function is 0.94.
Source: Own elaborations on ECHP data.

3. Differential growth: beyond Gibrat?

Let us consider the starting incomes $Y_{0,w}$ and $Y_{0,p}$ of two sets of individuals, both at the beginning of their working life, and let us assume that their initial income is respectively $Y_0^w=y_w$ (wealthy) and $Y_0^p=y_p$ (poor), where, obviously, $y_w>y_p$. At seniority *a*, *a* years later, following Gibrat, we expect their average incomes to be:

$$\begin{cases} E\left[\ln Y_a \mid \ln(Y_0 = y_p)\right] = \ln y_p + a \times \mu_R \\ E\left[\ln Y_a \mid \ln(Y_0 = y_w)\right] = \ln y_w + a \times \mu_R \end{cases} \tag{13}$$

where μ_R is the average of $\ln(R_i)$, as in eq. (5). In short, the average difference between the two groups should remain unchanged, *a* years later, and the same holds for the variance of the conditional distributions (see eq. 6). In other words, in Gibrat's model, once population homogeneity is assumed, no differential behavior should emerge between the poor and the rich.

But is this what really happens ? Let $\mathbf{y}_a$ be the vector of the incomes of a group of workers at seniority *a*, and ${}_a\mathbf{r}_b$ the vectors of their growth rates between seniority *a* and *b*, and let us consider the following regression

$$\ln {}_a\mathbf{r}_b = \beta_0 + \beta_1 \ln \mathbf{y}_0 + \boldsymbol{\varepsilon} \tag{14}$$

Gibrat's model implies that the regression parameter $\hat{\beta}_1$ should not differ significantly from 0. But the empirical estimates of Table 1 suggest, instead, that $\hat{\beta}_1 <0$, which means that small incomes increase faster than large ones.

Table 1 – *Estimation of equation (14) (Italy, 1995-2000).*

	Estimate	Std. Er.	t val.	Pr(>|t|)
$\hat{\beta}_0$	3.133	0.2776	11.29	<2e-16 ***
$\hat{\beta}_1$	-0.291	0.0276	-10.57	<2e-16 ***

Model: $\ln(Y_{2000}/Y_{1995}) \sim \ln(Y_{1995})$. Statistics: Adjusted R-squared: 0.1057; F-statistic: 111.8 on 1 and 936 DF, p-value: < 2.2e-16.
Source: own elaborations on ECHP data

This differential behavior between small and large incomes appears to be a general characteristic of income evolution, observed by several authors in different countries and epochs (Fields and others 2006; Salinari and De Santis 2008), especially where income inequality is stronger (e.g. in Latin America), and in periods of recession. Note, incidentally, that this is in part good news in terms of

the chances of getting out of poverty: the income of the poor tends to increase more rapidly than that of the whole population. The bad news is that this might be merely a "mechanical" consequence of the fact that incomes cannot be negative. Consider, as an illustrative example, two rich individuals who happen to have one a large increment and the other a large decrement of income. The following year, their average may be roughly the same, although variance is now high. But if we start with two poor individuals, the case of a large increment in the income of one of them cannot be offset by a large decrement in the income of the other, merely because the other can, at most, lose just the small income he used to earn. This point will be taken up again later, in Section 4.

Similarly, let us consider the following regression

$$\hat{\boldsymbol{\varepsilon}}^2 = \gamma_0 + \gamma_1 \ln \mathbf{y}_0 \tag{15}$$

where $\boldsymbol{\varepsilon}$ are the residuals of eq. (14). Once again, the idea is that, if Gibrat's model works well, the estimate of the regression parameter $\hat{\gamma}_1$ should not differ significantly from zero. But the empirical estimates of Table 2 indicate that the model does not adequately capture this aspect of reality: $\hat{\gamma}_1 < 0$, which implies that small incomes increase with a larger variance (in R) than large ones.[4]

Table 2 – *Estimation of equation (15) (Italy, 1995-2000).*

	Estimate	Std. Er.	t val.	Pr(>|t|)
$\hat{\gamma}_0$	0.438	0.1195	3.67	0.000259 ***
$\hat{\gamma}_1$	-0.030	0.0119	-3.34	0.000880 ***

Model: $\hat{\boldsymbol{\varepsilon}}^2_{2000} \sim \ln(Y_{1995})$; Statistics: Adjusted R-squared: 0.0107; F-statistic: 11.14 on 1 and 936 DF, p-value: 0.00088.
Source: own elaborations on ECHP data

4. Modifying Gibrat's model

In short, Gibrat's model, when used in a life-cycle perspective, proves satisfactory in describing several aspect of the general dynamic of personal income:

[4] We have also tried non parametric specifications, with moving averages and kernel. The results (not shown here) are qualitatively the same as in table 2, or better. Besides, we have also regressed on various intermediate years, that is $\hat{\boldsymbol{\varepsilon}}^2_{1999} \sim \ln(Y_{1995})$, $\hat{\boldsymbol{\varepsilon}}^2_{1998} \sim \ln(Y_{1995})$, etc., and we have tried to eliminate the top and bottom 5% of the distribution. Almost all of the results that we have obtained (not shown here) confirm the findings of tables 1 and 2.

a) the income distribution at seniority a;
b) the income distribution in year t;
c) the evolution of the mean and the variance of income distribution at different seniorities;
d) the autocorrelation (social mobility) between the income distribution at seniorities a and a+k;

Unfortunately, Gibrat's model is not fully satisfactory when it comes to describing the differential dynamic behavior of large and small incomes, in terms of both rates of growth and variance of growth. We will try now to find an explication for these findings.

As a first step, it is worth noticing that there are several ways of representing Gibrat's model. For example, eq. (1) can be rewritten as:

$$y_{a+1} = y_a + y_a(R-1) \tag{17}$$

That leads to the additive representation of the model

$$y_{a+1} = y_a + S \tag{18}$$

where S is a random variable, the average and variance of which are proportional to y_a, and y_a^2, respectively. This variable, S, can be thought of as the difference between two random variables: the gains *G* and the losses L that an individual can incur during a certain year (both with average and variance proportional to y_a, and y_a^2, respectively)

$$y_{a+1} = y_a + G - L \tag{19}$$

Equation (19) is a different, but equivalent way, of writing equation (1). Let us now modify a few of its underlying assumptions, so as to obtain a better fit with empirical data. Our extra assumptions now are that:

a) gains and losses are apportioned among individuals as in a Bernoulli experiment;
b) the probability of getting a unit of loss or a unit of gain is proportional to individual income;
c) an individual cannot lose more than its entire income.

Let T_Y, T_G and T_L stand for, respectively, the total of incomes, gains and losses observed in a certain cohort at seniority *a*. The probability of a unit of loss (or gain) is therefore $p = y_a/T_Y$. G can be conveniently represented with a Poisson distribution with parameter $\lambda_G = pT_G$, while the L-distribution will be a truncated Poisson distribution, with parameter $\lambda_L = pT_L$.

These extra assumptions produce two consequences that help understand the differential behavior of incomes starting from different levels, high and low (see Section 3):

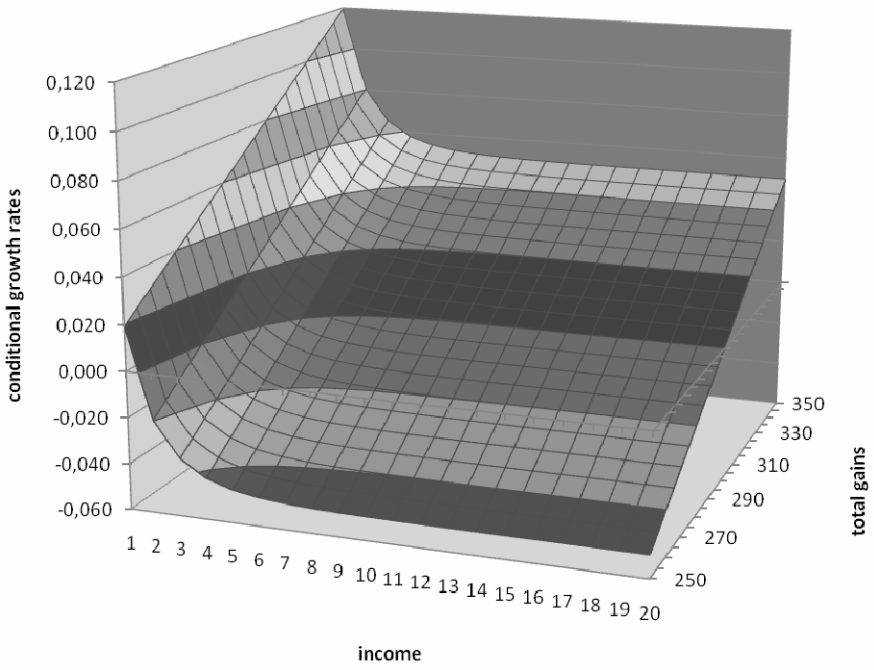
a) gains can span from 0 to infinity, but losses can only span from 0 to y_a. In other words, at least in theory, an individual can earn much more than the preceding year, but he/she cannot earn less than zero, which means that, at most, he/she can only lose all his/her income. Small incomes are comparatively advantaged in this respect, because they cannot suffer very large losses;
b) the variance of the two variables G and L increases, roughly, as a linear function of income (because in a Poisson distribution the mean and the variance are the same). This means that the variance of increments (R) decreases as a power low function of income (k/y_t).

In order to illustratively see how things work in practice, let us assume that total income is T_Y =1000, and that total losses and gains are T_L=T_G=300. We can now calculate the mean growth rate for an income y_a=5: the average gain will be 300*(5/1000)=1.5, while the average loss will be [300*(5/1000)]/F_P(5)=1.48, where F_p(5) is the value of the Poisson repartition function at point 5. The average increase will then be (1.5-1.48)/5=0.004. With a larger income, y_a=10, and the same procedure, we get an average increase of 0.0002, that is almost twenty times as small as that of the preceding case.

The same approach can be applied to the conditional variance of the growth rates: in the case of y_a=5 the variance of the gains is 1.5 (the same as the mean), while the variance of the losses is 1.4 (we obtained this value by previously calculating the density function of the truncated Poisson distribution with parameter 5/1000). The conditional variance of the growth rate will then be (1.5+1.4)/5^2=0.116. In the case of y_a=10 the conditional variance of growth rates is 0.060, that is almost twice as small as in the previous case.

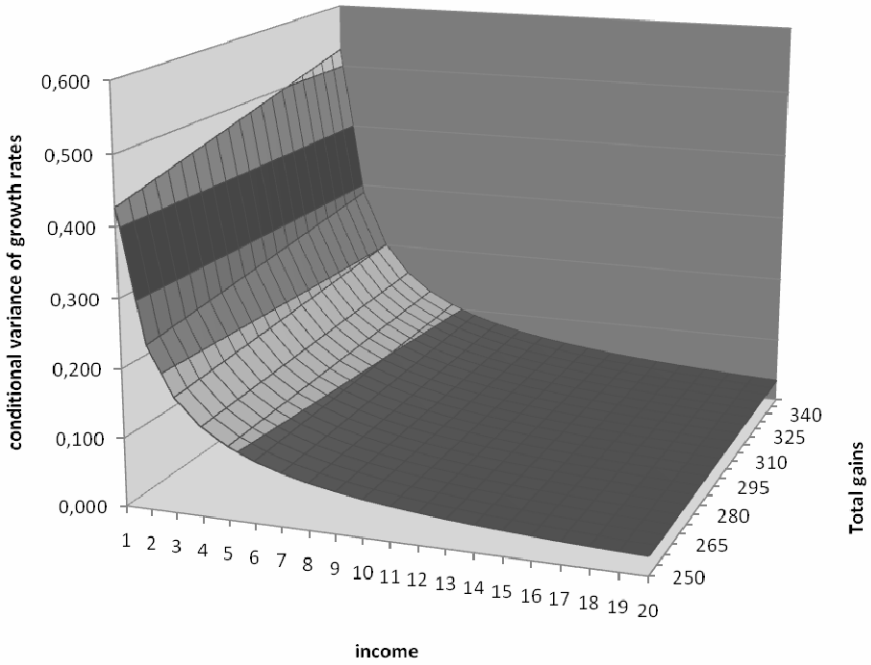
Figures 6 and 7 show the result of these exercises when the total gain G assumes values ranging from 250 and 350, and the individual income y_a ranges between 1 and 20.

Figure 6 – *Simulated expected growth rates of the modified Gibrat's model.*



Note: We assume initially that the gains *G* are distributed as a Poisson distribution, and that the losses *L* are distributed as a truncated Poisson distribution in the interval 0-y_a. Total income T_Y is set equal to 1000 and total losses T_L to 300. Then we calculate the mean increment for different income values y_a=1,2,…,20 and for different total gains value G_a =250, 255,…,350.

Figure 7 – *Simulated variance of growth rates of the modified Gibrat's model.*



Note: We assume initially that the gains *G* are distributed as a Poisson distribution, and that the losses *L* are distributed as a truncated Poisson distribution in the interval 0-y_a. Total income T_Y is set equal to 1000 and total losses T_L to 300. Then we calculate the mean increment for different income values y_a=1,2,…,20 and for different total gains value G_a =250, 255,…,350..

5. Conclusions

Gibrat's model is one of the few general theories on the process of income evolution, and therefore also income inequality and poverty. This paper shows that Gibrat's model describes satisfactorily the evolution of incomes along the life cycle, and that most of the critics that the model received in the past derive in part from misinterpretation, and in part from the lack of truly longitudinal data. An application to panel data suggests that the major theoretical (dynamic) consequences that the model predicts do correspond to what happens empirically.

We also found, however, that Gibrat's model does not explain the differential evolution of large and small incomes, in terms of both level and variance of the rates of growth. An improvement in this respect may come from the use of the so-called "modified Gibrat's model", which modifies only marginally Gibrat's assumptions. For average or above-average income, both models (Gibrat and modified Gibrat) produce basically the same results, because the effect of the truncation are negligible. But significant differences emerge when smaller income are considered. Further analysis is needed, however, to determine the formal characteristics of this (partially new) stochastic process.

Gibrat's model, both in its original and in its modified version, describes the evolution of incomes in a theoretical, homogeneous population, where no differential behavior exists among subgroups, but where, nonetheless, income inequality eventually emerges, as well as poverty. But both models can be also applied to the description of a heterogeneous population: in this case, a different set of parameters will characterize the various subpopulations. And if the parameters that drive the evolution of income vary systematically between population subgroups (e.g. the educated and the non educated; men and women; north and south; etc.), or with the social policies of certain countries or governments, then a parsimonious way of describing the genesis of income inequality, and poverty, will have be identified.

Acknowledgements

Financial support from the MIUR (Italian Ministry of Education, University and Research -Research grant 2007S8JLFN_001) and from the EU (6th Framework Programme; Contract No. CIT5-028571 - MAGGIE - Major Ageing and Gender Issues in Europe) is gratefully acknowledged.

Appendix.1 Covariance between incomes at seniority a and a+k (eq. 12)

$$\operatorname{cov}\left[\ln Y_a, \ln Y_{a+k}\right] = E\left[(\ln Y_a - \mu_a)(\ln Y_{a+k} - \mu_{a+k})\right] =$$
$$= E\left[(\ln Y_a - \mu_a)(\ln Y_a + \sum_{i=a+1}^{k} \ln R_i - \mu_a rk) \right] =$$

$$= E\left[(\ln Y_a - \mu_a)(\ln Y_a + \sum_{i=a+1}^{k} \ln R_i - \mu_a rk) \right] =$$

$$= E\left[\left(\sum_{i=0}^{a} \ln R_i - ra \right) \left(\sum_{i=0}^{a+k} \ln R_i - r(a+k) \right) \right] =$$

$$= E\left[\left(\sum_{i=0}^{a} \ln R_i - ra \right) \left((\sum_{i=0}^{a} \ln R_i - r(a)) + ((\sum_{i=a+1}^{k} \ln R_i - k(a)) \right) \right] =$$

$$= E\left[\left(\sum_{i=0}^{a} \ln R_i - ra \right)^2 + \left(\sum_{i=0}^{a} \ln R_i - ra \right) \left(\sum_{i=a+1}^{k} \ln R_i - k(a)) \right) \right] =$$

$$= E\left[\left(\sum_{i=0}^{a} \ln R_i - ra \right)^2 \right] + E\left[\left(\sum_{i=0}^{a} \ln R_i - ra \right) \left(\sum_{i=a+1}^{k} \ln R_i - k(a)) \right) \right] =$$

$$= Var[Y_a] + Cov\left[Y_a, \sum_{i=a+1}^{k} \ln R_i \right] = a\sigma^2 + 0 = a\sigma^2$$

Appendix.2 Correlation between incomes at seniority a and a+k (eq. 12)

$$\rho^2 = \frac{a^2\sigma^4}{a\sigma^2(a+k)\sigma^2} = \frac{a}{a+k}$$

$$\rho = \sqrt{\frac{a}{a+k}}$$

References

Armatte M. 1995. *Robert Gibrat et la loi de l'effet proportionnel*, Mathémathiques et Sciences Humaines, Vol. 129, pp. 5-35.

Atkinson A.B. 1997. *Bringing Income Distribution in from the Cold*, The Economic Journal, Vol. 17, No. 441, pp. 297-321.

Champernowne D.G. 1953. *A Model of Income Distribution*, The Economic Journal, Vol 63, No. 250, pp. 318-351.

Davies J.B. e Shorrocks A.F. 2000. *The Distribution of Wealth*, in Atkinson A.B. e Bourguignon F. (eds), Handbook of Income Distribution, Vol. 1, Chapter 11, pp. 600-675, Elsevier.
Fields G.S., Duval Hernández R.D., Freije Rodríguez S.F. e Sánchez Puerta M.L. 2006. *Income Mobility in Latin America*,
http://www.cid.harvard.edu/Economia/Mexico06%20Files/fieldsetal%20102306.pdf.
Gabaix X. 1999. *Zipf's Law for Cities: An Explanation*, The Quarterly Journal of Economics, Vol. 114, No. 3, pp. 739-767.
Gibrat R. 1931, *Les inégalités économiques: applications aux inégalité des richesses, à la concentration des entreprises, aux populations des villes, aux statistiques de familles, etc., d'une loi nouvelle, la loi de l'effet proportionnel*, Librerie du Recueil Sirey, Paris.
Gottschalk P. 1997. *Inequality, Income Growth, and Mobility: The Basic Facts*, The Journal of Economic Perspective, Vol. 11, No. 2, pp. 21-40.
Gottschalk P. e Smeeding T.M. 2000. *Empirical Evidence on Income Inequality in Industrial Countries*, in Atkinson A.B. e Bourguignon F. (eds), Handbook of Income Distribution, Vol. 1, Chapter 5, pp. 261-307, Elsevier.
Hart P.E. 1976. *The Comparative Statics and Dynamics of Income Distributions*, Journal of the Royal Statistical Society. Series A (General), Vol. 139, No. 1, pp. 108-125.
Jenkins S.P. 2000. *Modelling Household Income Dynamics*, Journal of Population Economics, Vol. 13, No. 4, pp. 529-567.
Kalecki M. 1945. *On the Gibrat Distribution*, Econometrica, Vol. 13, No. 2, pp. 161-170.
Mitzenmacher M. 2003. *A Brief History of Generative Models for Power Law and Lognormal Distributions*, Internet Mathematics, Vol. 1, No. 2, pp. 226-251 (http://www.internetmathematics.org/volumes/1/2/pp226_251.pdf).
Neal D., Rosen S. 2000. *Theories of the Distribution of Earnings*, in Atkinson A.B. e Bourguignon F. (eds), Handbook of Income Distribution, Elsevier, Vol. 1, Chapter 7, pp. 379-427.
Rutherford R. S.G. 1955. *Income Distributions: A New Model*, Econometrica, Vol. 23, No. 3, pp. 277-294.
Salinari G., De Santis G (2008), On the Evolution of Household Income, Luxembourg Income Study Working Paper Series, No. 488, July 2008.
Shorrocks A.F. 1976. *Income Mobility and the Markov Assumption*, The Economic Journal, Vol. 86, No. 343, pp. 566-578.
Sutton J. 1997. *Gibrat's Legacy*, Journal of Economic Literature, Vol. 35, No.1, pp. 40-59.

SUMMARY

Gibrat's model (1931) describes the evolution of income better than originally thought: panel data, now that they start to become available, are indeed log-normally distributed (as predicted by the model), while cross-sectional distributions are mixtures of log-normal curves. Gibrat's only weakness emerges in the prediction of the evolution of small incomes, but this can be remedied by a relatively minor change in one of Gibrat's assumptions. This leading to what may be called a "modified Gibrat's model", the characteristics of which are rapidly discussed in the paper.

Giambattista SALINARI, Dept. of Statistics, University of Florence.
Gustavo DE SANTIS, Dept. of Statistics, University of Florence.

INFORMAZIONI GENERALI

1. La collaborazione alla Rivista è aperta soltanto ai Soci della Società Italiana di Economia Demografia e Statistica e a coloro che vengono invitati a collaborare dalla Direzione.

2. I lavori per la pubblicazione vanno presentati alla Direzione della Rivista, adottando le regole indicate di seguito, con l'indicazione dei rispettivi titoli accademici ed eventuali qualifiche.

3. I lavori inviati verranno sottoposti al vaglio di un Comitato di lettura. Quelli non accettati per la pubblicazione non si restituiscono.

4. ***Gli autori degli articoli e delle note pubblicate rinunciano, in favore della Rivista, alla proprietà letteraria.*** La riproduzione integrale o parziale di articoli e note pubblicate nella Rivista è consentita alla condizione che risulti citata la fonte. Gli autori ricevono, gratuitamente, 15 estratti. Le spese per tirature supplementari o per stampa anticipata degli estratti sono a totale carico degli Autori.

5. La Rivista è distribuita ai Soci in regola col pagamento della quota di associazione.

6. La Rivista pubblica recensioni o segnalazioni delle opere ricevute in omaggio o in cambio.

Le quote di adesione alla S.I.E.D.S. sono le seguenti:
Soci individuali: Euro 70 *Soci collettivi:* Euro 120
Soci vitalizi, secondo accordi

Il versamento della quota sociali può essere effettuato tramite:
* c/c postale n. 42223008 intestato a Società Italiana di Economia Demografia e Statistica – Roma;

* bonifico bancario sul c/c **n. 000400852787** presso Unicredit Banca di Roma AG. 153 Città Universitaria, Piazzale Aldo Moro, 5 – 00184 Roma;

CODICE IBAN: IT97 D 03002 03371 000400852787

INFORMAZIONI PER GLI AUTORI

REGOLE PER LA COMPOSIZIONE DEI TESTI

Foglio formato A4. Spazio riservato al testo rettangolo di 13x18 cm. Impostazione margini: Superiore 5,8 cm; Inferiore 5,8 cm; Sinistro 4 cm; Destro 4 cm.

Il **titolo** deve essere scritto in **Times New Roman 12 punti**
Il **testo** deve essere scritto in **Times New Roman 11 punti**
Interlinea singola o 0,5 cm o 12 punti
Rientro prima riga (**capoverso**) **0,5 cm**
La numerazione delle pagine deve essere fatta fuori del rettangolo 13x18.

Testo, note, riferimenti bibliografici e appendici in Times New Roman 10 punti
Interlinea **delle note e dei riferimenti bibliografici 10 punti.**

PRIMA PAGINA

Nella prima pagina va indicato il titolo del lavoro (massimo 3 righe) lasciando una linea bianca, in carattere Times New Roman 12 punti, tutto maiuscolo, in grassetto centrato.

Il Nome e Cognome dell'autore va scritto in carattere Times New Roman 11 punti (come il testo), sotto il titolo lasciando una linea bianca, maiuscolo e centrato.

Il testo deve iniziare a cm 5 dalla prima riga del rettangolo di cm 13x18 e deve essere **giustificato.**

PARAGRAFI

I paragrafi devono essere numerati progressivamente con numeri arabi seguiti dal punto e dal titolo. Sia il numero, sia il titolo devono essere in grassetto e allineati a sinistra. Esempio:

1. Introduzione
2. Metodi statistici

I sottoparagrafi devono essere numerati progressivamente all'interno del paragrafo, sempre in numeri arabi e il titolo in corsivo. Analogamente al paragrafo vanno allineati a sinistra. Esempio:

3.1 *Descrizioni dati*
3.2 *Strategie di modellazione*

Sia i paragrafi, che i sottoparagrafi devono essere preceduti da due linee bianche e seguiti da una linea bianca; dopo il paragrafo (sotto-paragrafo) si inizia con un rientro (capoverso). Per le ulteriori numerazioni si procede in modo analogo.

FORMULE

Le formule devono essere numerate progressivamente con numeri arabi tra parentesi. Il numero deve essere allineato a destra. Il richiamo avviene nel testo tramite il numero della formula. Esempio: "…come espresso nella (7)…"

NOTE

Le note al testo devono essere numerate consecutivamente, ridotte al minimo, e riportate a piè di pagina. Il numero della nota nel teso va scritto in esponente con carattere più piccolo. Esempio: "…come affermato[1] …"

TABELLE

Le tabelle devono essere numerate consecutivamente con numeri arabi. Le tabelle devono essere contenute nel formato 13x18 cm, usando anche caratteri più piccoli. L'intestazione inizia con **Tabella n** (in grassetto) allineata a sinistra e il titolo è separato da un trattino e scritto in corsivo. Esempio:
Tabella 2 – *Stranieri residenti in Italia nei censimenti del 1991 e del 2001.*

FIGURE E GRAFICI

Le figure devono essere numerate consecutivamente con numeri arabi. Le figure e i grafici devono essere stampati in ***bianco nero e non a colori.*** L'intestazione inizia con Figura n (in grassetto) allineata a sinistra e il titolo è separato da un trattino e scritto in corsivo. Esempio:
Figura 1 – *Funzione integrata di rischio per maschi e femmine.*

RIFERIMENTI BIBLIOGRAFICI

Le citazioni bibliografiche nel testo si effettuano con il nome dell'autore o degli autori e l'anno. Esempio: Cicchitelli, Herzel e Montanari (1992), oppure (Cicchitelli, Herzel, Montanari, 1992).

Dopo le conclusioni seguono i Riferimenti bibliografici (in grassetto) e allineato a sinistra. I riferimenti devono essere in ordine alfabetico come segue: il cognome precede il nome puntato e l'anno di pubblicazione. Il titolo del lavoro deve essere scritto in corsivo. Esempio:

Riferimenti bibliografici

Barr S.R. e Tuner J.S. 1990. *Qualità Issues and Evidence Statistical File*. In: Lieping G.E. e Uppuluri V.R.R. (eds), Data Quality Control. Theory and Pragmatics, pp 245-313, Marcel Dekker, New York.

Cicchitelli G., Herzel A. e Montanari G.E. 1992. *Il campionamento statistico*. Il Mulino, Bologna.

Trivellato U. 1990. *Modelli di comportamento e problemi di misura nelle scienze sociali: alcune riflessioni*, Atti della XXXV Riunione Scientifica della S.I.S., vol. 1, pp 11-31, Cedam, Padova.

APPENDICI

Dopo i riferimenti bibliografici seguono le appendici (lasciare soltanto due righe bianche dai riferimenti bibliografici) che devono essere numerate con numeri arabi. L'intestazione inizia con Appendice n (in grassetto) allineata a sinistra e il titolo è separato da un trattino e scritto in corsivo. Esempio:
Appendice 1 – *Eventuale titolo*
Appendice 2 – *Eventuale titolo*

SUMMARY

In fondo al testo segue il riassunto in inglese (summary) che deve essere contenuto in 1 pagina del formato 13x18 cm. Il titolo, in carattere 11 punti, deve essere scritto minuscolo, in grassetto e centrato.

TITOLI ACCADEMICI E ALTRO

Indicare in fondo al testo i titoli accademici o altro al momento della presentazione del testo.

Consigli utili:

- per evitare inutili sprechi nell'impaginazione definitiva del volume della rivista non inserire pagine o spazi bianchi;
- usare il carattere **grassetto neretto** con parsimonia;
- evitare il più possibile le maiuscole, nelle sigle di enti o altro usare il carattere MAIUSCOLETTO;
- per le parole straniere inserite nel testo usare il carattere corsivo;
- usare le virgolette soltanto quando si riportano brani originali;
- per gli elenchi puntati, i sottoelenchi ecc. non inserire il rientro del capoverso (0,5 cm), usare eventualmente i diversi tipi di elenchi numerati messi a disposizione dal programma;
- nella riquadratura delle tabelle inserire semplicemente quei bordi orizzontali ($_{1/4}$ pt) che servono per una migliore lettura della tabella ed evitare i bordi verticali e troppo marcati;
- inserire tabelle, grafici e immagini dopo il punto di fine capoverso;
- spostare in fondo al testo tabelle, grafici e immagini che non rientrano nel formato e necessitano quindi di una riduzione fotografica, segnalando l'inserimento a tutta pagina nel testo.

INVIARE IL TESTO IN LINGUAGGIO WORD PER WINDOWS ALL'INDIRIZZO E-MAIL DI VOLTA IN VOLTA INDICATO

Esempio: lasciare una linea bianca

STUDI EMPIRICI SULLE PICCOLE E MEDIE IMPRESE ITALIANE

Mario Rossi

_____Prima riga del testo (a cm 5 del formato 13x18)_____

SOCIETÀ E RIVISTA ADERENTI AL SISTEMA ISDS
ISSN ASSEGNATO: 0035-6832

Direttore Responsabile: Prof. ENRICO DEL COLLE

Iscrizione della Rivista al Tribunale di Roma del 5 dicembre 1950 N. 1864

Associazione all'Unione Stampa Periodica Italiana

TRIMESTRALE
Spedizione in abbonamento postale art. 2 comma 20/c legge 662/96 - Filiale di Roma

---

La copertina è stata realizzata da Pardini, Apostoli, Maggi p.a.m.@tin.it - Roma