Firm Size Distribution: Testing the "Independent Submarkets Model" in the Italian Motor Insurance Industry*

Luigi Buzzacchi, Politecnico di Torino
and
Tommaso Valletti, London School of Economics and CEPR**

Discussion Paper
No. EI/24
September 1999

The Toyota Centre
Suntory and Toyota International Centres
for Economics and Related Disciplines
London School of Economics
Houghton Street
London WC2A 2AE
Tel: 020 7955 7719

---

*We thank John Sutton for his helpful comments.
**Corresponding author: Department of Economics, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, Tel: +44 020 7955 7551, fax: +44 020 7831 1840, email: t.valletti@lse.ac.uk

Abstract

This paper tests the presence of multiple independent submarkets in the Italian motor insurance industry. Independence is motivated by administrative boundaries among provinces and by further locational reasons. We find that the independence effects are sufficient to induce a minimum degree of inequality in the size distribution of firms once submarkets are aggregated. These results are fully consistent with the predictions of Sutton (1998). We also show that the degree of skewness in the firm size distribution is related to characteristics such as the population living in an area, its density and the riskiness of a submarket.

Keyword: size distribution of firms, independent submarkets, insurance companies
JEL Classifications: L11, G22

1. Introduction

The notion of "market definition" is crucial to any work in Industrial Economics. When economists try to understand the way firms interact, with the aim of giving theoretical answers to either positive or normative questions (or both), they would start by positing demand functions that define the relevant market(s) where firms compete. This would also be true of course for empirical work (possibly leaving the task to define a market to the data).

In principle, one could try to split an industry into smaller and smaller subindustries, until a break in the chain of substitutes is found. In this way, one would be left with a certain group of subindustries each containing groups of products with a low degree of substitutability between groups and a high degree of substitutability within groups. Taking this argument to its extreme and logical conclusion, the question arises as to what kind of effects one could expect to find once an industry is made of many submarkets, all independent of each other. In particular, should we expect to find in a market a certain number of identical firms, each of whom has entered roughly the same submarkets? Or, conversely, should we expect to find skewness in the size distribution of firms because there are bigger firms that have entered relatively more submarkets?

These questions lie at the core of a sizeable literature on the growth of firms that started long ago with the work of Robert Gibrat. Gibrat proposed a "law of proportionate effect", stating that the rate of growth of a firm, relatively to its market size, is a random variable, independent of the firm's current size. Thus bigger firms should grow faster and the firm size distribution should follow a log-normal distribution. Later research used more sophisticated stochastic models in order to try to generalise Gibrat's results. Yet in the end, no distribution or families of distributions could be used across industries with some degree of general validity (see Sutton, 1997, for a review of the literature).

The question of the appearance of skewed distributions of firm size, is posited again in a new form and answered in Sutton (1998). Sutton does not look for "the" family of distributions that data should fit, rather he looks for a "bound" on the form of size distribution. The problem is addressed by developing a theory of submarkets that are independent in the sense that each firm's profit function is additively separable into the contributions of the firm's profits in different submarkets. A high number of submarkets means that there are many entry opportunities to be filled by firms, some of whom may be operating in many other submarkets, and some of whom may enter the industry for the very first time. Sutton shows that when the number of submarkets is sufficiently large, strategic considerations become less important than statistical effects. In particular, by introducing a simple "Symmetry Principle", that specifies that there is an equal treatment with respect to all potential entrants to each submarket, Sutton finds that the size distribution of firms in the industry (at the aggregate level) has to be greater than the following bound:

(1) $C_k \quad k/N\,[1 - \ln(k/N)]$

where C_k is the k-firm concentration ratio and N is the total number of firms in a market. Hence there is a minimum degree of inequality in firm size that, in graphical terms, implies that the Lorenz curve constructed with data for a certain market has to be more skewed than the limiting value given by the right-hand side of eq. (1). Sutton thus is able to find a bound on the firm size distribution, giving an important theoretical contribution to the growth-of-firms literature.

While aggregation effects operate over many submarkets and give some predictions on the structure one could expect to find in the overall market, strategic effects are predominant within each submarket. From a theoretical point of view, very little could be said on the structure of individual submarkets. However, under special circumstances, some restrictions can be placed on the size distribution. Sutton (1998) shows that a maximally fragmented structure in which all firms have the same size should arise in models with symmetric product differentiation when two conditions are satisfied:

- Price competition is weak;
- Products are close substitutes.

These models are particularly appropriate when there are overlapping submarkets, for instance where firms are supplying customers in different geographical regions. These elements, that predict maximal fragmentation at the submarket level and skewness in aggregate, do coexist, for instance, in the US cement industry analysed by Sutton (1998).[1]

The purpose of this paper is to test the “independent submarkets model” in the Italian motor insurance industry, which represents an ideal setting. In the first place, there are administrative boundaries such that every vehicle has to be insured in the province where the owner is resident. Since there are 103 provinces in Italy, this immediately gives us an idea of the minimal amount of submarkets that are present. Clearly, provinces are of very different sizes, hence we cannot expect every province to contain the same number of submarkets. On the other hand, demographics show a big variance over provinces and we can try to understand which provinces should contain more or less submarkets.

[1] The bound on the inequality of firm size distribution has been recently tested by other authors. Walsh and Whelan (1999) consider the Irish food and drink industry, where submarkets are defined in terms of consumers’ tastes rather than geographical location. The locational dimension, on the other hand, is crucial in de Juan (1999), who uses data from the Spanish retail banking market, and whose work has many parallels with our own. See also Asplund (1999) for a related analysis of the Swedish driving schools market, where strategic effects are prevalent in each submarket.

Secondly, in Italy the majority of premiums are collected via tied agents, while other sales structure are of negligible importance. Tied agents dominate the distribution system in the non-life segment, and in particular in the motor business where the share of tied sales outlets is estimated to be well above 90%. The number of agencies in each province therefore provides excellent information on the points of presence of a company. Some companies have a strong regional presence, but most have branches all over the country, resulting in overlapping submarkets. Maximal fragmentation is thus expected at the submarket level as long as the two conditions that we mentioned before are satisfied.

Products are indeed homogeneous (this is definitely the case in the motor liability segment that we analyse, accounting for about 50% of total non-life volume in Italy) and price competition is limited. This last remark deserves further comment. Until 1 July 1994, the Italian insurance industry was subject to substantive government regulation that manifested itself chiefly in fixing price ceilings and was especially noticeable in the field of motor liability insurance. The government's intention in fixing price ceilings was to place insurance regulation in the service of its social and economic objectives (above all the fight against inflation). After deregulation occurred, the industry has not witnessed the increase in competition envisaged by policy makers. Both the industry regulator and the Antitrust authority have repeatedly intervened in the market. We do not know whether firms are colluding or if there are simply some "teething" problems after many decades of direct regulation; however, the evidence is that firms, even now, do not advertise their products, their price lists are not published, quotes on premiums can be obtained only upon request, and an upward trend in prices has even been observed over the last five years.

In the section that follows we give a description of the industry (more data can be found in the Appendix) and we find that fragmentation is observed in submarkets whereas their aggregation produces a Lorenz curve that lies above the theoretical bound. We also show that the degree of skewness in the firm size distribution is related to characteristics such as the population living in an area, its density, and the riskiness of a submarket. In the final section, we briefly summarise our findings and argue that the skewness in the firm size distribution that we have found is not just an Italian phenomenon but it is also observed in many European countries.

2. Firm size distribution in the Italian motor insurance industry

The insurance industry is typically divided into two quite separate businesses: Life and Non-Life. The former includes a variety of pension schemes and investment activities, while the latter concerns motor insurance, accident and health, fire, theft, general liability and credit. Their weight in the Italian insurance business is relatively similar: for instance in 1998, total premiums were 111,278 billion lira (approximately £ 40 billion), 50.8% collected in Life insurance and 49.2% in the Non-Life segment.

In this paper we concentrate on the Non-Life business: in particular, the Motor Insurance segment is the class that represents the bulk of direct premiums collected (58% in 1998). Finally, the Motor Liability class has the major influence on the Motor segment (82% in 1998; and 48% of the entire Non-Life business). In Italy, the legal system imposes the obligation of drawing up certain insurance policies and the most wide-spread compulsory insurance is by far Motor Liability: every vehicle has to be insured (in general, it is the vehicle and not the owner as in other countries).

What makes this especially interesting for our purposes is that a vehicle has to be insured within a particular administrative area. Italy is divided into 20 Regions and 103 Provinces: vehicles are registered in a province, hence a policy has to be bought in the same province. Since the majority of vehicles are privately-owned cars, the province of residence of a (potential) car owner is an excellent proxy for the maximum dimension of an independent submarket. Car-owners can be perfectly discriminated on the basis of their residence and they must insure according to the policy designed by a company for that particular province.

Within a province, the motor insurance industry has a further and very strong spatial dimension.[2] Such a spatial nature thus reinforces the first, broad cut deriving from the administrative boundaries and the distribution system becomes fundamental to explain the size distribution of insurance companies. In Italy, the absolute preponderance of premiums is written through agents (89% of total Life and Non-Life) and the preponderance is even more striking in the Non-Life business (in the Life segment financial advisors and bank counters are increasingly important). We do not have a figure for premiums collected in the Motor Liability class through agents, however the companies that use exclusive agents collect 97.5% of total premiums and they rely almost solely on such a distribution system. It can even be argued that the number of agents is a better measure of presence in a submarket than premiums, since the latter would also include money collected using other systems that do not really belong to the same market.

[2] Insuring a car is an activity that typically occurs once a year and a contract is tacitly renewed unless a six-month notice is given. Typically, other (and repeated) contacts are expected should an accident occur, thus increasing "transportation" costs. Moreover, information among consumers is very disperse: there is virtually no advertising done by companies, and prices can be collected only by directly contacting a company. Sales over the phone or over the Internet are still extremely small in Italy (see the Appendix). We recall that the limited intensity of price competition is suggested by the numerous interventions of the Italian Competition Authority over the past few years. Motor Liability prices have also been regulated until 1994 and many customers are still not very aware of the potential gains that could be made by searching for better tariffs within a province. As a result, a customer is likely to patronise a company that operates in the closest neighbourhood of the area where he lives and/or works: any other place, even in the same province, would be too far away and the unknown price would not justify a longer trip.

We have collected data for all tied agents of 66 companies present in Italy in 1998 in every province (see the Appendix for details). The number of agencies that each insurance company has in a given territory gives an indication of its points of presence and it is the proxy we use for market shares. Every time an “opportunity” opens in some area, a new agency is also likely to be opened. The independence between opportunities in the Motor Liability segment, both spatial and administrative, provides an ideal setting in which to empirically examine the validity of Sutton’s independent submarkets model.

The aggregation process is clear: figure 1 shows the industry Lorenz curve constructed using the number of agencies each company has. Such a curve satisfies Sutton’s bound given by eq. (1). We have also drawn a second Lorenz curve, using total premiums rather than agencies. As it can be seen, the two curves are not different from each other.[3]

[Figure 1 – Lorenz curves in the Italian motor insurance industry (national level)]

While the limiting distribution is clearly confirmed in aggregate, our aim is now to see whether a more fragmented result holds within each submarket. As we said before, in Italy there are at least 103 well-defined submarkets corresponding to the provinces. Figure 2 reports the Lorenz curves for 103 provinces: the majority of the Lorenz curves lies below Sutton’s bound. This is already a satisfactory result, and one could simply argue that in some provinces strategic effects are more important than in other provinces. However, we do not find this interpretation very appealing; on the contrary, we think that analysing each province in greater depth, statistical rather than strategic effects should still prevail.

[Figure 2 – Lorenz curves for the Italian motor insurance industry (province level)]

Provinces are very different in their demographics and physical characteristics: population varies from 90,000 to almost 4 million; surface vary from 212 square km to 7,520 square km. It is implausible to assume that each province represents a single submarket, thus some aggregation process should occur also at the province level. It is immediate to think that as population grows,

[3] The degree of skewness in firm size distribution can be described numerically using the Gini coefficent (see below for its definition). In figure 1, Gini is 58.96% if premiums are used and 61.93% if agencies are used (as a reference point, the Gini corresponding to Sutton’s bound is 49.97%). We performed the Kolmogorov-Smirnov test to see whehter the two alternative measures could produce significant differences in the degree of skewness and we rejected with a confidence above 99% the hypothesis that the two curves are drawn from different distributions. Using agencies rather than premiums could have an impact on the skewness of a Lorenz curve, especially once it is drawn for a submarket (for instance, think of scale effects). We do not have data on premiums collected by firms within a given area. However

more profitable opportunities should arise, causing more skewness in the Lorenz curve (population is strongly correlated with the number of vehicles, r = .93; see the Appendix). This is indeed the case: by adopting a cut-off level of 800,000 inhabitants, the number of curves above the bound in figure 2 would be very small. However, even in that case, there would be some significant outliers, including for instance a couple of densely populated but small provinces. Our aim in the remaining part of this section, is to investigate further what could determine the number of submarkets within each province and to test whether fragmentation emerges as the number of submarkets decreases.

The first experiment we carried out was as follows. Each province has an administrative capital, which is typically the biggest town in the province (there are only 2 exceptions out of 103). If we assume that people living in the capital town subscribe to policies in the same town (which is likely), then we do not have to worry too much about the extension of the province. Figure 3 draws the Lorenz curves constructed at the level of each capital town, excluding the 10 biggest cities (with population above 300,000 people). Almost all the Lorenz curves lie below the limiting curve, supporting our hypothesis that towns are rather well-defined submarkets, with smaller towns containing fewer opportunities. Of course there are also other effects: for instance people living in towns other than the capital could well be working in the biggest town and buy their insurance there, hence the population outside the main town could help explain the size distribution of the main town itself.

[Figure 3 – Lorenz curves for the Italian motor insurance industry (capital town level)]

We have also constructed Lorenz curves for the agencies outside the provincial capital. These curves are not reported here because of space limitations, but they prove to be systematically more skewed than the curves for the capitals. There are of course differences between populations, but we also believe that – for a fixed population - more submarkets should arise when the population is dispersed rather than concentrated in the same place. To test this idea we start using a standard index of firm inequality that is positively related to the degree of skewness in firm size, the Gini coefficient. The Gini coefficient for a population of x_i observations, $i = 1, \ldots, n$ with average $\bar{x}$ is defined as follows:

$$Gini = \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\left|x_i - x_j\right|}{2\bar{x}n^2}$$

in the Appendix we perform some tests that reassure that the two series, premiums and agencies, are not statistically different. To the extent that premiums and agencies are not proportional, our results below need to be qualified.

We have calculated Gini coefficients for every province, both at the aggregate level and after having split the province in its capital and the rest. If we plot the pairs Gini-outside capital with Gini-capital within the same province, we obtain a cluster of points that lie around the 45 line (see figure 4, panel a). What is more interesting is to divide the whole sample into 2 subgroups: provinces that have populations roughly equally distributed in the capital and outside (panel b) as opposed to provinces where there are more people living outside the provincial capital (panel c). In the latter group, with relatively more people outside the provincial capital - and living in a more dispersed way - we argue that more independent opportunities should arise, and hence more aggregation effects outside the capital might be expected. This is confirmed in the figure: the Gini-extra-capital is systematically higher than the Gini-capital in the subsample (panel c), while the two coefficients are relatively similar where population is more balanced between capital and outskirts. The comparison of the pairs of Gini coefficients within the same province is particularly interesting because we can reduce the importance of other influential factors (such as income, risk, etc.) that can be easily assumed to be similar within the same province.

[Figure 4 – Comparison of the Gini coefficients in the capital town
and outside the capital town of the same province]

The last exercise relating to the possible determinants of the numerosity of opportunities – thereby causing different levels of aggregation - has to do with the riskiness of the motor insurance industry. The phenomenon of fraud is well known (for example, in 1996 companies reported thousands of cases of bogus or suspicious claims; in the same year ascertained fraud entailed a payment of 2.7% of total claims incurred by firms in the Motor Liability segment). One parameter that could be used is the frequency of accidents that are 10% of insured vehicles on average but can be as low as 6% in some provinces and almost 20% in other provinces. Concentrating on capital towns only – in order to get rid of other spatial effects – we find a positive relationship between Gini coefficients and frequency of accidents (figure 5, panel a). However, this result should not suggest that markets with more accidents present more opportunities! On the contrary, the positive relationship depends on size effects: bigger towns have more vehicles and also more opportunities, and more accidents are likely to occur in an area where lots of vehicles are circulating than in a deserted area. In towns with a similar number of vehicles, however, we could expect a similar probability of occurrence of an accident: differences in probabilities could be an indicator of a market that is riskier, either for adverse selection reasons (drivers are less skilled) or for moral hazard reasons (fraud). In figure 5, panel b, we have isolated the smallest towns: they clearly represent the lower-left end of the total sample and, among them, there is no positive relationship between Gini and frequency: if anything, the correlation is even negative, giving some mild support to the idea that, after controlling for size effects, more accidents occur in riskier markets that are also less attractive.

[Figure 5 – Frequency of accidents and Gini coefficients in capital towns]

We do not have a model that could be tested in a more rigorous way. However, we have performed some very simple regressions that confirm in quite a striking way what we have argued before. We have regressed the Gini coefficient of a submarket, calculated according to three different levels of aggregation (whole province, capital towns only, whole province without capital towns), against various variables. The results for some regressions are reported in table I. They all investigate the same idea: the more opportunities arise, the more important aggregation effects should become, producing higher degrees of skewness and higher Gini coefficients.

Most of the variance of the Gini coefficient of a certain submarket is explained by the population in the same submarket. This is reasonable: the greater the number of people, the more opportunities arise in that submarket. The same result can be obtained by putting vehicles in place of population, since they are strongly correlated with the population. It is important to stress that we have regressed Gini against the *logarithm* of the population (in table I, just compare the increase in the explanatory power of regression A *vs*. B at the province level; an even stronger effect would result by using total premiums rather than population). This indicates that opportunities do not arise linearly with population: once a new person appears, that individual will represent a new opportunity for low levels of population (assuming that at first, people will be scattered all over the place), but eventually a new person will appear in the same place where other people have already appeared, and this does not give the same rise in opportunities. In other words, 10 cars to be insured in 10 different places sufficiently distant from each other may support 10 independent opportunities, while the same 10 cars at the very same location represent fewer opportunities (in the limit they could even belong to only one independent submarket). This finding is reinforced by the fact that people are not distributed uniformly over a territory, rather they aggregate in some places. An alternative interpretation is that people own vehicles of different values: a few expensive vehicles may support the same number of opportunities as many cheap ones, but this effect is captured in the variable premiums per capita (see also below for a discussion).

Regression C at the province level also includes premiums per capita. We should first remark that this regression has the same explanatory power as a regression that uses the logarithm of total premiums only (not reported in the table). This is hardly surprising, but since we do not have total premiums at a more disaggregated level, we have reported such a regression for better comparisons with regressions at the level of capital towns and outside them.[4] Finally, regression D shows that the frequency of accidents and a geographic variable (surface) are not statistically significant at the province level.

In the regressions at the levels of capital towns and outside them, we have also included the residual population, to indicate that there could be interaction between capital towns and areas outside them in the presence of people's mobility (think of commuters). The results show a significant positive effect on the Gini of the capital towns of the population living outside and not viceversa. This makes sense, because there are many reasons why some categories of people regularly go to the capital town (which is typically the most important town), where they can subscribe their policies, thus giving rise to new opportunities, while the reverse is less likely.

Outside capital towns, we also tested whether a geographic dimension of the market such as its surface could help to explain the aggregation process. We tested this with a slight twist, including a dummy (DLD) that takes a value of 1 in areas with low population density and 0 otherwise (in particular, low density is defined as less than 150 inhabitants per square kilometre; using this threshold, 60% of the provincial areas outside the capital town would be defined as "low density"). Results indicate that surface has a positive and significant impact on Gini only in areas outside the capital towns that are not densely populated (see regression B'). This is coherent with the fact that the product in question has strong features of horizontal differentiation: given an amount of vehicles to be insured, there are more independent opportunities when they distributed over a wide territory. However, there is a "scale" effect: once density is high enough, then the surface effect is not relevant anymore and population alone is the main explanatory variable. This is in line with what we said before on "urbanisation", i.e. the tendency of people to aggregate in some specific places in a province.

[4] We cannot regress Gini coefficients over *both* the log of population and the log of premiums per capita at the province level. This would be equivalent to having in the regression the log of population and the log of total premiums that are highly correlated (r = .96). We also would like to stress that the variable premiums per capita represents a measure of the attractiveness of a market. Premiums per capita are in fact higher in richer provinces. In Italy, richer provinces are typically in the North and in the Centre, as opposed to poorer Southern regions. The latter are also typically the riskier markets, in the sense that more cases of fraud are reported by companies. To support our last claim, premiums per capita are negatively correlated with the frequency of accidents, r = -.23, and positively correlated with the average cost of an accident, r = .59 (bear in mind that the most expensive claims for companies are those that involve the loss of human lives, while cases of fraud are typically concerned with less expensive accidents).

	PROVINCE				*CAPITAL TOWN*	*OUTSIDE CAPITAL TOWN*	
REGRESSION	A	B	C	D		A’	B’
CONSTANT	0.347 (36.68)**	-1.014 (9.89)**	-1.075 (11.91)**	-1.047 (10.45)**	-1.246 (14.00)**	-1.083 (7.34)**	-1.178 (9.09)**
PROVINCE POPULATION	0.104 (9.17)**						
LN(PROVINCE POPULATION)		0.110 (13.87)**	0.106 (15.30)**	0.103 (10.49)**			
LN(CAPITAL TOWN POPULATION)					0.093 (9.36)**	0.021 (1.45)	0.021 (1.48)
LN(POPULATION OUTSIDE CAPITAL TOWN)					0.034 (3.86)**	0.087 (6.46)**	0.092 (7.41)**
PREMIUMS PER CAPITA			0.308 (5.65)**	0.323 (5.24)**	0.205 (3.06)**	0.351 (3.42)**	0.388 (3.79)**
FREQUENCY OF ACCIDENTS				0.001 (0.15)	-0.001 (0.17)	-0.004 (0.70)	-0.002 (0.41)
SURFACE				0.003 (0.90)		0.005 (1.03)	
SURFACE*DLD							0.008 (1.98)*
R^2	45.42	65.58	73.92	74.15	80.95	59.00	60.17
ADJUSTED R^2	44.88	65.24	73.40	73.09	80.17	56.89	58.11
F	84.04	192.44	141.71	70.26	104.09	27.92	29.30

t-statistics in parenthesis
* Significant at the 5% level
** Significant at the 1% level

Table I – Firm inequality in submarkets: regression results

Finally, we have included in our regressions for capital towns and outside them also premiums per capita collected by companies and the frequency of accidents, with the intention of capturing a measure of the attractiveness and of the risk involved in a market. We did not find any impact deriving from the frequency of accidents, while premiums per capita are always positively significant. Premiums per capita are correlated with various measures of wealth and development of a province (higher premiums are asked for more expensive vehicles that are typically insured in wealthier areas, where there is also a higher number of vehicles per capita): higher premiums correspond to more opportunities, and the positive relationship with the degree of skewness is confirmed in all the regressions we performed.[5]

To summarise, the empirical evidence confirms that the aggregation process of different independent submarks can explain the skewness in the firm size distribution. Higher Gini coefficients are found as more opportunities arise, where the number of opportunities can be reasonably represented by total premiums or by the population size when data on premiums are not available. The logarithmic specification captures the idea that populations are not uniformly

[5] The lack of (negative) significant correlation between Gini coefficients and the riskiness of a market, as measured by the frequency of accidents, has also to be assessed against the presence of potential multicollinearity problems due to the correlation between the frequency of accidents and both dimensional and wealth variables. The Italian scenario is characterised by higher risks in big towns (r = .53 between frequency and population; this fact could also in part explain the logarithmic specification that we have used) and in “poorer” towns (typically in the South). With the data at our disposal it is not possible to separate neatly these effects.

distributed over a territory, rather there are relevant agglomeration features (there are more opportunities in 10 towns with a population of 100,000 inhabitants than in a single big city with 1 million people). The “spatial” dimension of submarkets is also further testified by a geographic variable (only below certain density levels) and by the explanatory power of the population living outside the capital town on the Gini of the capital town, capturing the very typical effect of flows of commuters. We did not find any impact deriving from adverse selection problems (captured by the frequency of accidents), while it is possible to argue that moral hazard problems have a negative impact on Gini, but we do not want to stretch our interpretation of the variable premiums per capita too far. Finally, it would be interesting to test, at the capital town level, the effect on the skewness of firm size distribution of “tougher” price competition, resulting from the simple fact that there is more neck-to-neck competition in a big town with many companies, than in a dispersed territory. However, we could not include the number of firms in our regressions for the very strong positive correlation it has with Gini, especially at the capital town level (see table 3 in the Appendix).

3. Concluding remarks

This paper has tested the role of independent submarkets as a determinant of firm size distribution, along the lines of the theory recently proposed by Sutton (1998). We have conducted the test in the Italian motor insurance industry, exploiting its interesting features:

- It is divided into 103 independent administrative areas;
- Within each area, location of companies is the most important parameter;
- Firms are typically present in multiple areas and do not seem to compete fiercely against each other.

We have found fragmented distributions at the submarket levels as opposed to very unequal distributions when submarkets are aggregated at the national level, thus offering very strong evidence in favour of the “bounds” approach. Moreover, we have investigated further what could determine the numerosity of independent opportunities that arise in each market.

The skewed distribution found at the national level does not seem to be a phenomenon limited to Italy. Using a different source of data (Swiss Re, 1996), we have information on concentration indices of insurance companies in the Non-Life business in different European countries. Since we do not have the total number of companies, we cannot draw the aggregate Lorenz curve for each market. However, we have data for the 5-firm, 10-firm and 15-firm concentration indices (measured on premiums), so we can perform in every country the conditional test proposed by Sutton (1998). According to the theory, for a fixed k m and a certain value C_m

for the m-firm concentration ratio, the actual k-firm concentration ratio should lie above the following bound:

(2) $C_k \quad k/N_m\, [1 - \ln(k/N_m)]$

where N_m is the solution of:

(3) $C_m = m/N_m\, [1 - \ln(m/N_m)]$

Figure 6 plots a scatter diagram of C_5 versus C_{15}. The solid curve corresponds to the fragmented equilibrium (slope = 1/3), while the dotted bound has been derived solving recursively eq. (3) with respect to N_m using $m = 15$ and a grid of values of C_m between 0 and 1, and then plugging every resulting value of N_m into the expression for the bound (eq. (2) with $k = 5$). All the data points lie very neatly above the bound. We have also repeated the exercise with C_{10} versus C_{15} (not reported here), always confirming the conditional prediction.

We have not addressed the problem of firm dynamics at all in this paper, despite the fact that it is central to the size distribution of firms. It would be very interesting to see what happens when independent markets have reached different levels of maturity and development (i.e., young versus old markets). In particular, the question of whether firms that are present at the national level are also older firms or whether companies with a strong regional vocation remain such over time or try to capture opportunities elsewhere, is a valuable one. These are questions that we plan to investigate in the future.

[Figure 6 – Testing the conditional prediction at the European level]

Bibliography

ANIA, 1999, *Italian Insurance in 1998 – Annual Report*, Rome

Asplund, Marcus, 1999, “On the Size Distribution of Firms and Markets,” mimeo, Stockholm School of Economics

ISVAP, 1997, Circolare 3495, Rome

de Juan, Rebecca, 1999, “The Independent Submarkets Model: An Application to the Spanish Retailing Banking Market,” mimeo, Fundacion Empresa Publica, Madrid

Sutton, John, 1997, “Gibrat’s Legacy,” *Journal of Economic Literature* 35: 40-59

Sutton, John, 1998, *Technology and Market Structure*, MIT Press, Cambridge (MA)

Swiss Re, 1996, “Deregulation and Liberalization of Market Access: the European Insurance Industry on the Threshold of a New Era in Competition,” *sigma*, n. 7

Walsh, Patrick Paul and Ciara Whelan, 1999, “Modelling Firm Size Distribution in Food and Drink Products,” mimeo, Katholieke Universiteit Leuven

Appendix

The Italian Motor Insurance Industry: firms included in the sample.

The empirical work in this paper is based on a database that comprises the geographic distribution system of 61 companies over 103 provinces. In total, 15,044 tied agencies were identified using two different sources: the biggest companies report their distribution system on their web sites, while the remaining agencies were collected using the telephone directories of all the Italian provinces. Agencies within each province were further subdivided according to whether they were or not in the capital town of the province. Using some random samples, we also checked that the two sources did not produce significant discrepancies (the telephone directories of 1998 report information collected at the end of 1997, while websites are usually updated at various points in time – in our case between the end of 1998 and early 1999).

In 1998 there were 97 insurance companies authorised to offer contracts in the Italian Motor insurance business. These companies collected total premiums amounting to 27,897 billion lira (22,932 for Motor Liability and 4,964 for other damages). Premiums are mainly collected through tied agents: this phenomenon is true at the European level, and exceptionally high in Italy (Swiss Re, 1996). Alternative sales systems include brokers, independent agents, direct sales, and distribution channels that are destined to become increasingly important such as banks, phone sales and Internet sales. Out the total population of 97 companies, in 1998, 6 companies sold only over the phone and over the Internet (0.6% of total premiums), 4 companies used only bank branches (0.1%) and 66 sold almost exclusively via tied agents (97.4% of total premiums). The remaining 21 companies sold exclusively through the remaining channels (independent agents, brokers, direct sales).

We concentrate our study only on those companies selling via tied agents, both because it is the dominant channel and because it is the channel that exhibits the clearest locational dimension. While phone sales and sales through bank branches compete against tied agents, this would not be true for the other distribution systems that are devoted to satisfy companies' needs (think of brokers, re-insurance – i.e. the way companies pool major risks). Overall the exclusion of firms not relying on tied agents does not seem particularly problematic.

Out of the remaining 66 companies, 5 had particular agreements such that the same distribution system was shared with other companies that belonged to the same group. In these cases we have merged the firms and considered them as a single entry. As a result our relevant population is made of 61 companies.

Summary statistics.

Table 1 contains the descriptive statistics relative to the 103 Italian provinces. Additional information has been taken from the Italian Insurers Association (Ania, 1999). Table 2 reports the correlation matrix between the exogenous variables involved.

It should be noticed that the Gini coefficient is on average below the Gini coefficient associated to Sutton's bound (0.5), but there are provinces with values above it, to indicate that aggregation effects are already working at the provincial level. Moreover, most of the companies attempt to pick opportunities all over the country since, on average, in every province half of the total number of companies are present with at least one agency.

	Mean	Standard deviation	Min	Median	Max
Number of branches	146	160	27	106	1,181
# companies (province)	33	6	18	33	56
# companies (capital town)	30	7	17	30	54
# companies (outside capital town)	22	8	2	23	42
Premiums collected (L. billion)	197	251	20	128	1,664
Herfindhal (province)	5.47%	1.49%	3.68%	5.11%	14.29%
Herfindhal (capital town)	5.37%	1.46%	3.50%	5.04%	16.50%
Herfindhal (outside capital town)	9.45%	7.42%	4.11%	7.06%	50.00%
Gini coefficient (province)	0.41	0.10	0.14	0.40	0.63
Gini coefficient (capital town)	0.30	0.11	0.10	0.30	0.64
Gini coefficient (outside capital town)	0.35	0.11	0.00	0.37	0.60
Population in the province (million)	0.56	0.62	0.09	0.38	3.80
Population in the capital town (million)	0.17	0.32	0.02	0.09	2.65
Population outside the capital town (million)	0.39	0.36	0.03	0.28	2.43
Surface (sq. km.)	2,925	1,750	212	2,562	7,520
Provincial density (inhabitants per sq. km.)	244	334	37	173	2,669
Provincial density outside the capital town (inhabitants per sq. km.)	163	211	26	130	1,775
Inhabitants per branch (thousand)	3.87	1.32	2.16	3.39	8.03
Premiums per capita (L. million)	0.34	0.09	0.15	0.37	0.52
Motor vehicles (million)	0.09	0.10	0.00	0.06	0.72
Frequency of accidents (number of accidents/vehicles insured)	10.10%	2.01%	6.36%	9.69%	18.24%
Average accident s cost (L. million)	4.48	1.10	2.29	4.40	8.11

Table 1. Summary statistics for 103 Italian provinces

Capital town pop.	Province pop.	Pop. outside capital	Premiums collected	Motor vehicles	Branches	Premiums per capita	Surface	Density (outside capital)	Frequency of accidents	Average accident s cost	
1.00	0.89	0.64	0.90	0.81	0.87	0.15	0.16	0.45	0.49	0.08	Capital town pop.
	1.00	0.92	0.96	0.93	0.93	0.09	0.23	0.69	0.53	0.09	Province pop.
		1.00	0.83	0.86	0.81	0.01	0.25	0.78	0.47	0.09	Pop. outside capital
			1.00	0.96	0.98	0.29	0.19	0.58	0.39	0.03	Premiums collected
				1.00	0.97	0.28	0.23	0.57	0.33	0.04	Motor vehicles
					1.00	0.28	0.22	0.54	0.33	0.02	Branches
						1.00	0.17	0.05	0.23	0.59	Premiums per capita
							1.00	0.24	0.06	0.16	Surface
								1.00	0.47	0.02	Density (outside capital)
									1.00	0.49	Frequency of accidents
										1.00	Average accident s cost

Table 2. Correlation matrix

There is a very strong correlation of Gini coefficients measured according to alternative definitions of submarkets. The correlation matrix is not reported here for the sake of brevity (however, see figure 4 in the main text), on the other hand we believe it is of some interest to report the correlation matrix between various measures of market structure (table 3). There is a positive correlation between Gini and the number of firms, while the degree of skewness in size distribution is not related to a classic measure such as the Herfindhal index. This is consistent with what one could expect from the independent submarkets model that predicts a positive relationship between number of firms and their degree of skewness, while there is a non monotonic relationship between the latter and the Herfindhal index. When there are a few opportunities, in fact, there will be a few symmetric companies, resulting in high Herfindhal indices but low Gini. On the other hand, in big provinces, it is likely that many opportunities are to be filled by companies: if such opportunities are more or less independent, statistical aggregation should give a picture where many companies of very different sizes coexist (high Gini and Herfindhal indices).

Province			Capital town			Outside capital town			
Gini	# companies	Herfindhal	Gini	# companies	Herfindhal	Gini	# companies	Herfindhal	
1.00	0.70	-0.04	1.00	0.75	0.03	1.00	0.66	-0.43	Gini
	1.00	-0.03		1.00	0.03		1.00	-0.71	# companies
		1.00			1.00			1.00	Herfindhal

Table 3. Correlation matrix between Gini and concentration indices

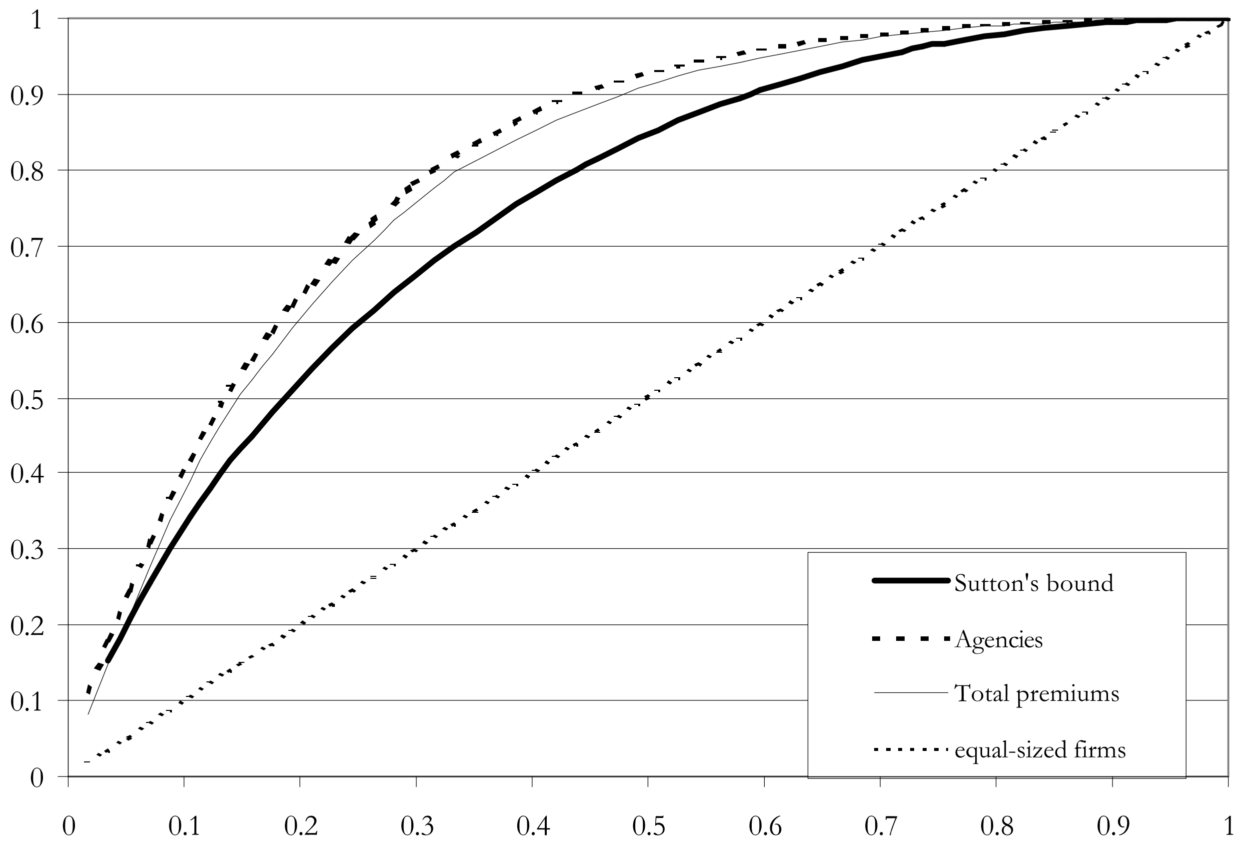
Agencies versus premiums.

The empirical analysis conducted in this paper is based on companies' market shares in a certain area, measured by the share of agencies that a company has in the same area. This choice is motivated by several reasons we have already mentioned and that are repeated here:

- Agencies represent by far the most adopted distribution system;
- In every submarket, a company typically collects premiums using other sales channels (brokers and re-insurance): these premiums are not related to the motor insurance business object of our study;
- According to Sutton's analysis, a firm grows by collecting opportunities in various submarkets: there is a logical correspondence in our case with firms' growth via the opening a new agency.

One could still be worried that the analysis based on agencies may be biased, for instance because it does not capture effects such as multiproduct purchases or economies of scope between various life and non-life segments. This is not really a counter argument as long as agencies collect premiums on top of motor liability in a way proportional to premiums collected in the motor liability segment. In this case, in fact, the Lorenz curves drawn with premiums and with agencies would be coincident. Unfortunately, using premiums rather than agencies would constitute a big problem in the empirical analysis because firms typically do not give information on premiums collected at the provincial level. However, we could carry some additional analysis based on a limited sample where total premiums were available at the provincial level (data collected from the Industry Regulator, ISVAP, 1997; for the national level see figure 1 and footnote 3). Based on 103 observations collected in 1998, we found a simple Pearson correlation of .943 (p-value <.0001) between total agencies and total premiums at the provincial level. Similarly, based on 61 observations, we found a simple Pearson correlation of .962 (p-value < .0001) between agencies and premiums at the company national level. Using the same series, we also tested that the cumulate distributions were not significantly different (using the Kolmogorov-Sminornov test), and finally we verified that rankings were preserved (using the Spearman test). These results give support to the idea that agencies are a good proxy for premiums collected.

Figure 1 - Lorenz curves in the Italian motor insurance industry
(national level)



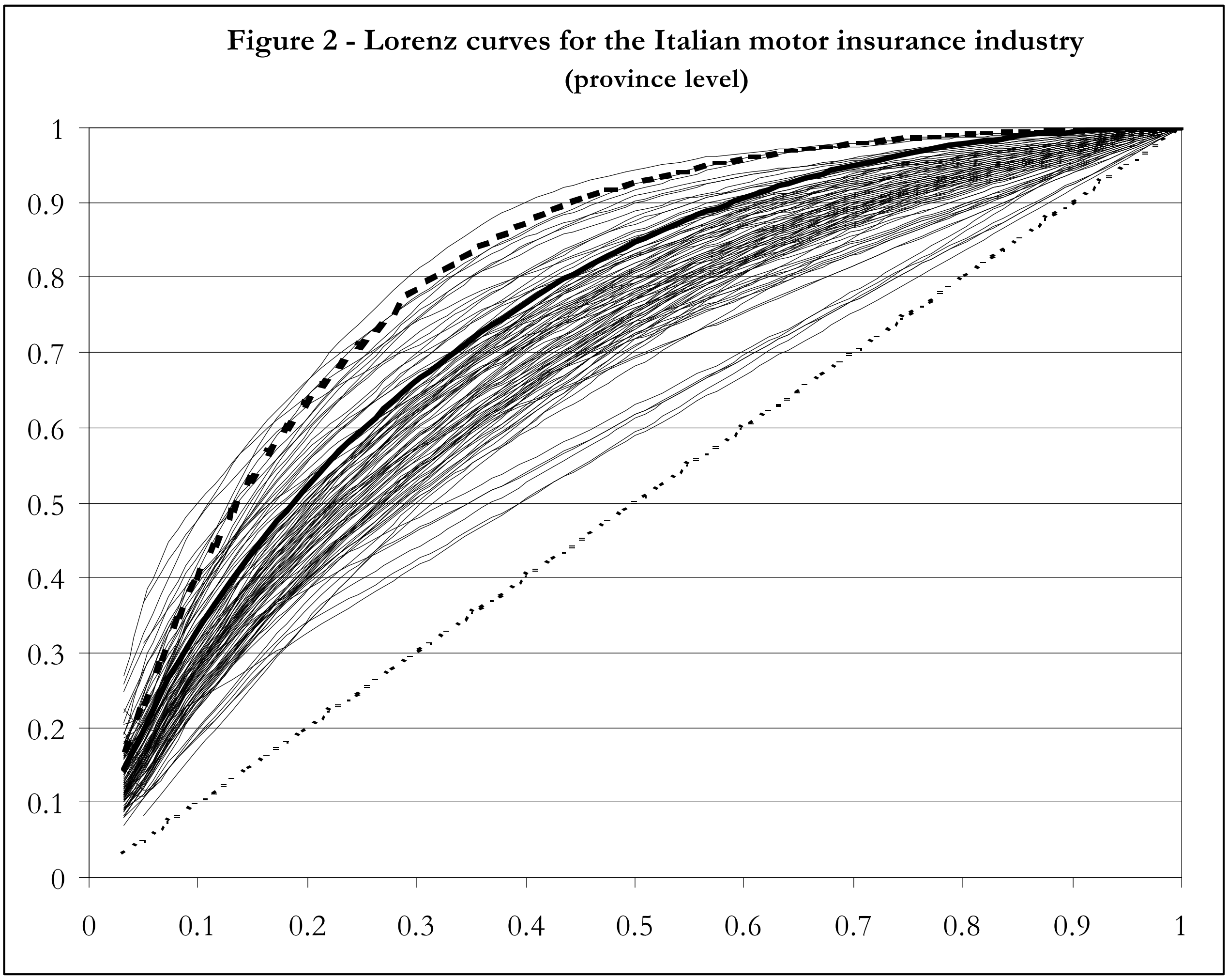


Figure 2 - Lorenz curves for the Italian motor insurance industry
(province level)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

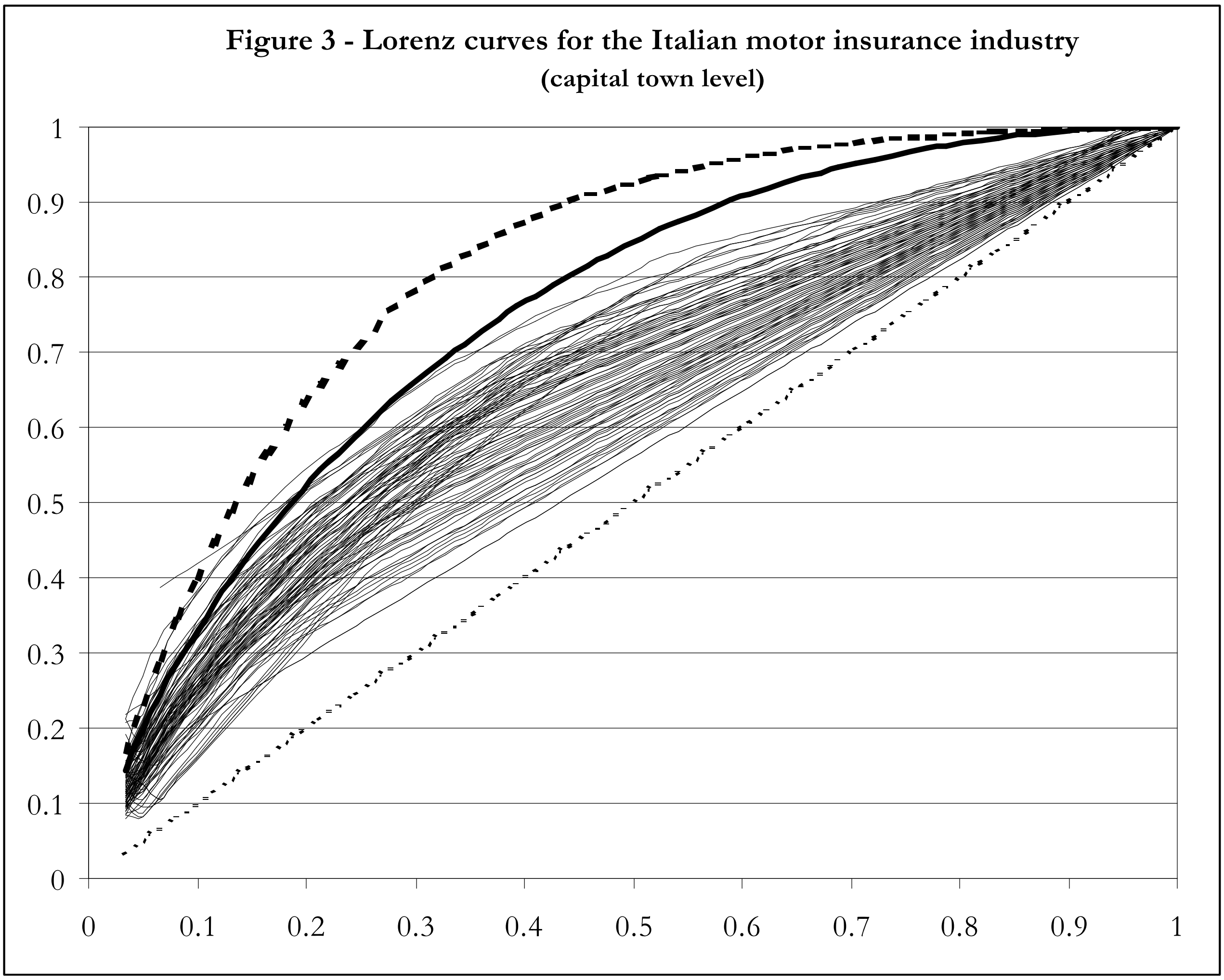


Figure 3 - Lorenz curves for the Italian motor insurance industry
(capital town level)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

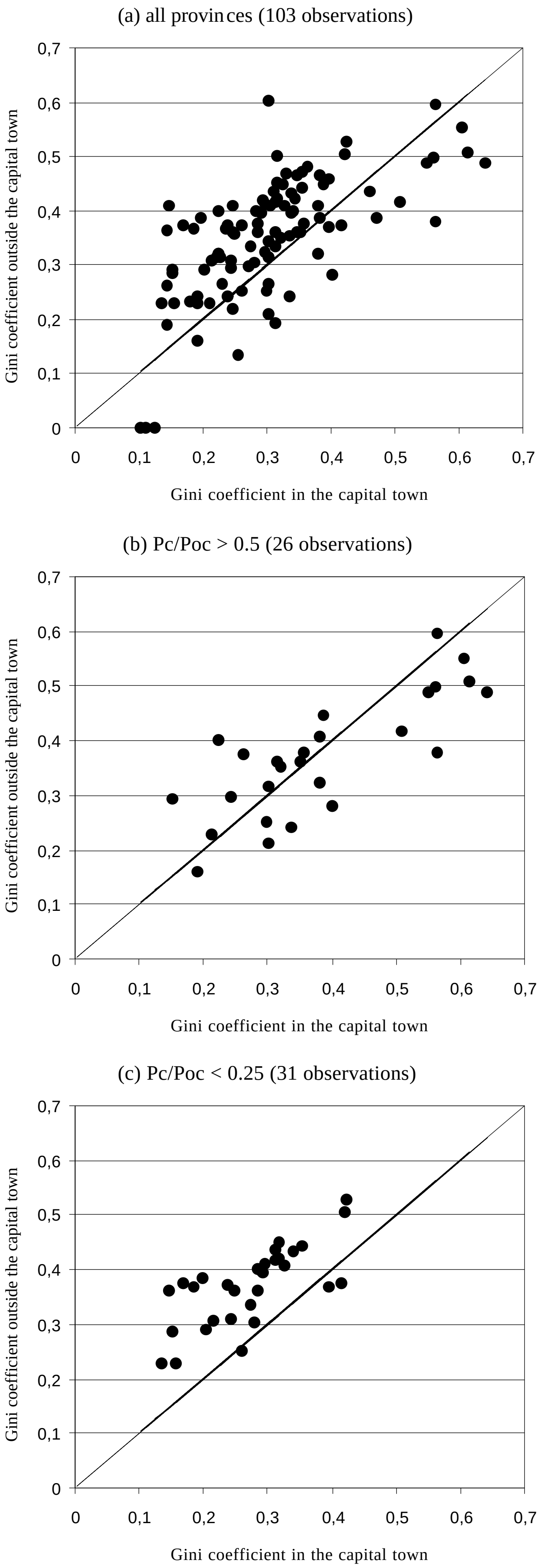


Figure 4 – Comparison of the Gini coefficients in the capital town and outside the capital town (Pc/Poc is the ratio at the province level between the population in the capital town and the population outside the capital town)

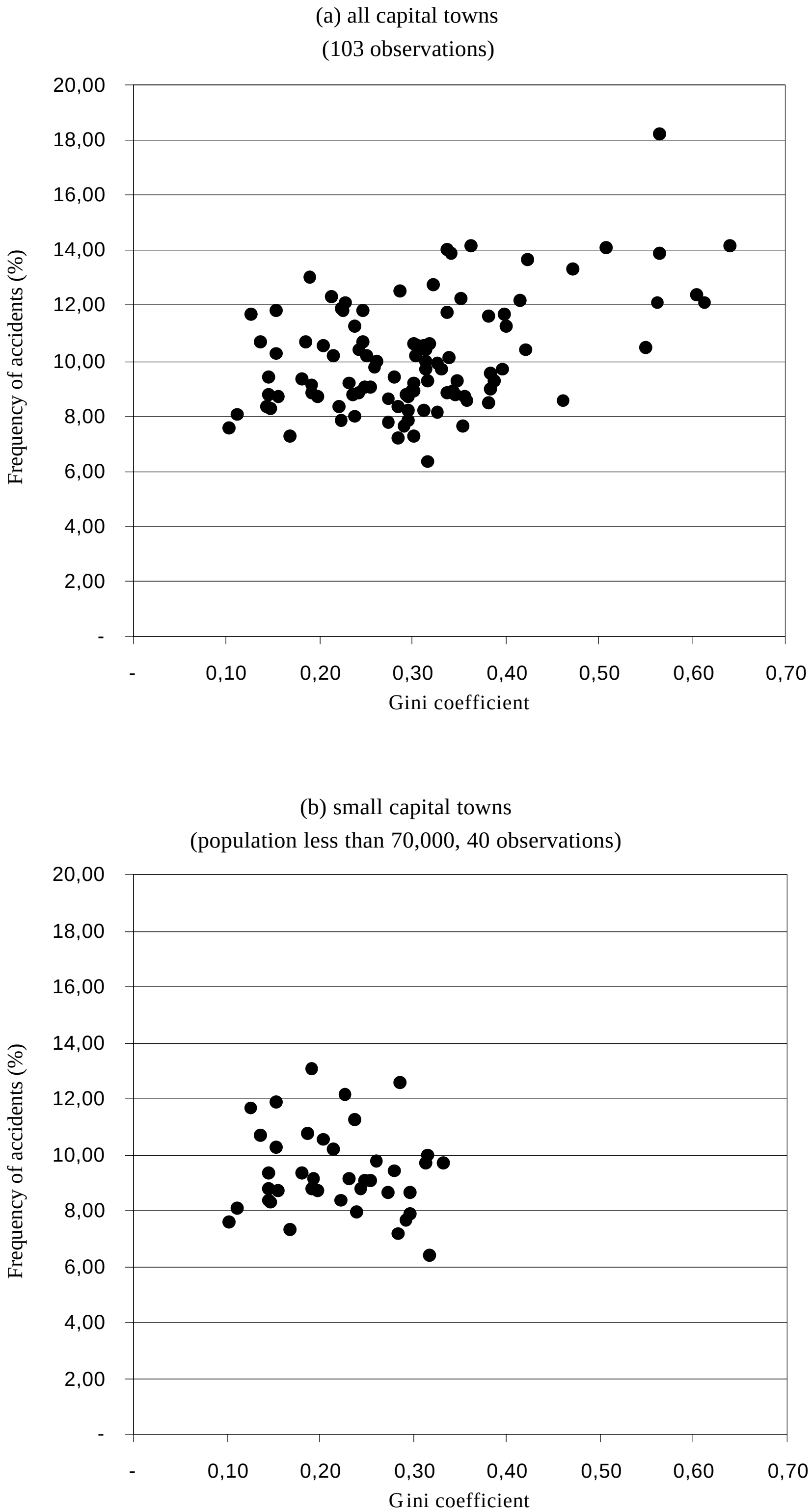


Figure 5 – Frequency of accidents and Gini coefficients in capital towns

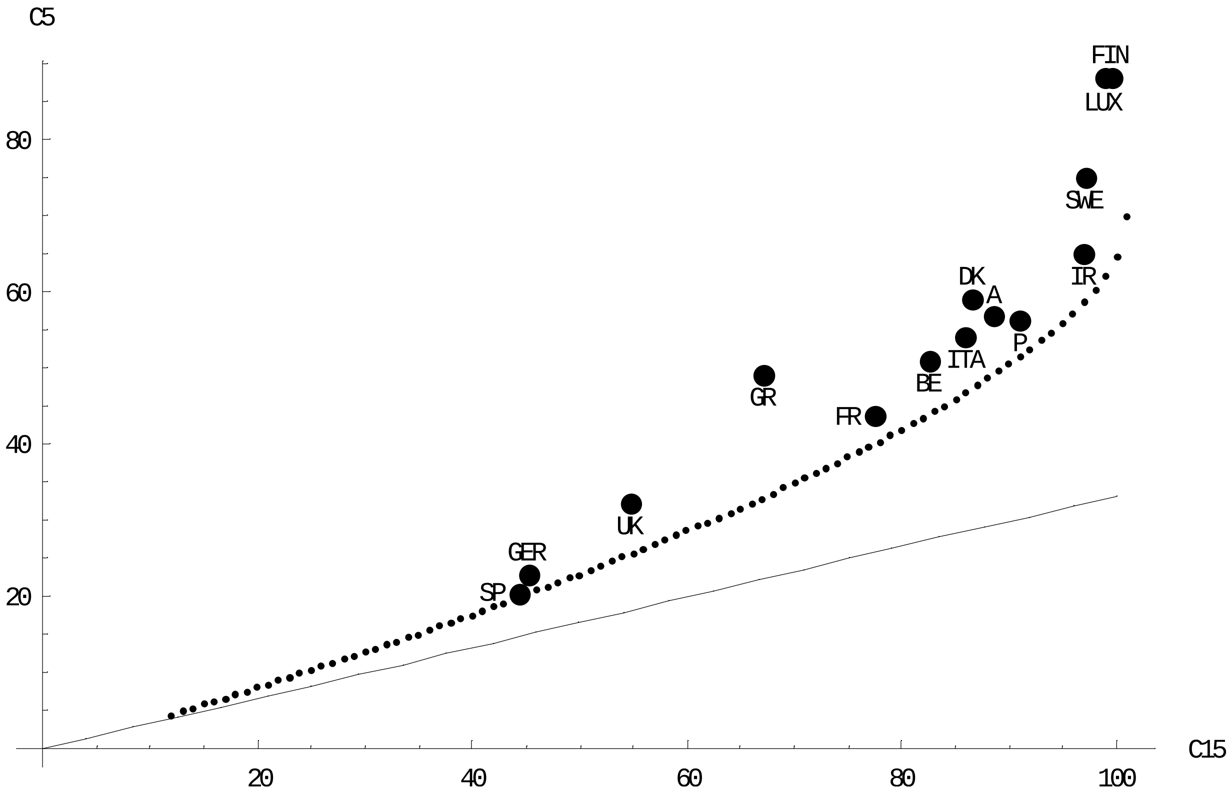


Figure 6 – Testing the conditional prediction at the European level