

Warsaw School of Economics

Collegium of Economic Analysis

mgr Kamil Gala

Abstract of PhD thesis titled

**Spatial aspects of insurance risk in automobile insurance**

PhD thesis

prepared under the guidance of

dr hab. Wojciech Bijak, prof. SGH

Warsaw 2020

## 1. Subject, purpose and scope of work

The subject of this dissertation is the spatial distribution of insurance risk on the Polish motor insurance market, including motor third-party liability insurance for motor vehicle owners (MTPL) and motor own damage insurance (MOD or *autocasco*). In view of the ambiguity and complexity of the concept of insurance risk (Ronka-Chmielowiec, 2013) in this work it is defined as the risk associated with concluded insurance contracts, which takes the form of insurance accidents and related claims and benefits. This definition emphasizes the likelihood of an insurance accident (one or more) and the amount of loss resulting from it.

The literature on the subject (e.g. Taylor, 2001; Brouhns et al., 2002; Gala, 2017) indicates that insurance risk (measured e.g. by claim frequency) shows noticeable spatial heterogeneity. The main goal of the dissertation is to assess the significance of the geographical dimension as a factor enabling better assessment of insurance risk in motor insurance. To this end, tools combining the experience of spatial econometrics and actuarial ratemaking have been developed and used.

*A priori* ratemaking is part of the insurance risk assessment process in which the premium is determined at the time the contract is concluded on the basis of the observable characteristics of the insured and the subject of insurance (e.g. vehicle or real estate). MTPL and MOD insurance constitute a significant part of the Polish property insurance market, which is due to the large number of motor vehicles in Poland and the obligation to have MTPL coverage of each vehicle. Therefore, statistical methods are the fundamental tool for determining insurance premiums in motor insurance.

Over the years, both specific actuarial methods (taking into account, for example, asymmetrical distributions of the dependent variables) and a set of factors considered to be good predictors of insurance losses (such as the age of the insured, the engine capacity of the vehicle or the region of its use) have been established. At the same time, spatial econometrics is developing largely without connection to actuarial statistics. These methods are used to study phenomena located in geographical space, most often characterized by spatial diversity and the relationship between processes occurring in neighboring areas (so-called spatial autocorrelation). One can indicate (see e.g. Lee and Pace, 2009) many real-world situations that lead to the observation of spatial autocorrelation - both as a result of real interactions between regions (so-called *contagion effect*), and as a result of measurement errors or omission of relevant variables of spatial nature. Especially the second effect may be significant in motor insurance, taking into account that the boundaries of administrative division units do not reflect

the specificity of road traffic and weather phenomena, which factors affect the risk of claim occurrence. Regardless of the actual source of autocorrelation, omitting it during the construction of the model may cause that the obtained parameter estimators will not have the desired properties, such as consistency, unbiasedness or efficiency.

The dissertation covers theoretical and empirical analysis of statistical models used for *a priori* ratemaking taking into account the spatial nature of insurance data. The first chapter is devoted to *a priori* ratemaking. The second chapter presents issues related to spatial data and describes the basic models used in spatial econometry. The third chapter formulates a generalized spatial linear model, then presents models with a multi-level factor as its specific case and discusses methods of spatial segmentation. In the last chapter, theoretical considerations were illustrated by empirical analysis of data on the Polish motor insurance market, which included a description of the spatial distribution of insurance risk, the estimation of parameters of predictive models and the assessment of the effectiveness of the methods and models under consideration.

## **2. Reasons for undertaking the research topic - its theoretical and practical significance**

In accordance with art. 33 of the Act of 11 September 2015 on insurance and reinsurance activity (Dz.U. 2020 poz. 895 as amended), the insurance undertaking is required to determine the the insurance premium after assessing the insurance risk (section 1). This premium should be set in an amount ensuring at least the fulfillment of all obligations resulting from insurance contracts and coverage of the costs of carrying out the activity of the insurance undertaking (section 2), and the undertaking collects relevant statistical data to determine the premium based on them (section 3).

Competition on the Polish motor insurance market, increasing customer awareness and legal regulations force insurance companies to improve their risk assessment methods. Consequently, there is a growing interest in both new data about the vehicle owner, such as his credit standing (Gala and Kolak, 2015), penalty points and traffic offenses (Pinquet et al., 2011), as well as data on the number of kilometers driven per year and the style of driving (Lemaire et al., 2016). It should be noted, however, that the use of these possibilities may be limited due to high costs, legal regulations or practical considerations (e.g. the need to obtain client's consent to access data from an external source). In this situation, the alternative may be to use more efficiently the data that is already in the possession of the insurance company. Identifying of

appropriate tariff variables and determining their impact allows not only to better match the premium to the risk associated with the insurance contract, but also to actively target the insurance undertaking's sales activities towards groups of clients with low claim frequency. It is also worth noting that, in a broader sense, the pricing policy of insurance companies may provide a financial incentive for specific behaviors that affect road safety (e.g. encourage drivers to choose cars with a lower engine power).

The importance of spatial aspects in the study of economic phenomena was emphasized in the work of Anselin (1988), who as the main issues indicates spatial heterogeneity and spatial autocorrelation. At the same time, he draws attention to the need to take into account spatial effects for the purpose of proper specification of the model, estimation of its parameters and verification by means of statistical tests. The author lists two main sources of spatial autocorrelation - errors related to the phenomenon measurement and data aggregation (territorial observation units do not reflect the spatial distribution of phenomena), as well as real processes related to the spread of phenomena and interactions between various areas in space (e.g. migration of people or transport of goods).

Both of these sources can be found in motor insurance. First of all, traditional tariffs at the level of a geographical unit (postal code, municipality, district) may be affected by an error resulting from the fact that the boundaries of these units often do not coincide with the occurrence of large urban agglomerations or communication routes (e.g. highways). Secondly, there may be interactions between different areas, e.g. in the form of heavy traffic between areas (e.g. commuting). In both cases, spatial autocorrelation will be observed at the level of cross-sectional data, which should be included in the specification of the tariff model. It is also worth paying attention to the fact that while more detailed spatial division could allow to reflect spatial nature of the studied phenomenon closely, the consequence of this approach is the small number of observations within territorial units or the lack of observations for some of them. In this situation, obtaining reliable results is difficult or even impossible.

There is extensive scientific literature both in the area of actuarial risk assessment in motor insurance and in the area of spatial data analysis. Issues related to actuarial risk assessment in motor insurance, both *a priori* and *a posteriori*, have been extensively described, including in the works of Denuit et al. (2007) and Lemaire (1995). In turn, issues related to spatial econometrics were developed, among others in the works of Anselin (1988) and LeSage and Pace (2001), while in Polish literature one can indicate a work edited by Suchecki (2010). It should be noted, however, that in actuarial literature there are relatively few papers in which the spatial aspects of insurance risk in motor insurance are analyzed in detail, and none of these