

Data-driven pricing

Opportunities in the non-life insurance market

The market competition in the field of Dutch non-life insurance is fierce. The ability to compare insurance policies online enables consumers to find the lowest prices for the products they need. This compels insurance companies to quote increasingly competitive premiums, whilst claims and operating costs show an increasing trend, resulting in decreasing profit margins. Therefore, now is the time for non-life insurers to reform their pricing strategies. We strongly believe a data-driven pricing strategy will become the new market standard, and have developed an open source-pricing platform to assist insurers in the transition.



drs. Rinze Valkenburg MSc AAG is a senior manager at KPMG Financial Risk Management.

valkenburg.rinze@kpmg.nl



Peter Bosschaart MSc is a senior consultant at KPMG Financial Risk Management.

bosschaart.peter@kpmg.nl

INTRODUCTION

The Dutch non-life insurance market is under pressure, and has been for a while. High combined ratios are common for several insurance products and a few big players dominate the market. However, market competition is fierce. The ‘old way’ of buying all your policies at the local insurance office has disappeared. Consumers are now more than ever able to compare insurance policies that cater to their needs, often causing them to search for the lowest available prices. This forces insurers to set competitive premiums, whilst managing claims and operational costs. These competitive premiums, combined with the trend of increasing claim levels and operational costs, result in decreasing margins with an even higher pressure to maintain a stable portfolio with profitable clients.

However, not only consumers have more possibilities to benefit from the increasing availability of information, also the insurance companies can profit from these. For example, insurance companies have more detailed insight in their consumer’s behavior, and the underlying risks, and are therefore able to transfer from traditional pricing strategies to more modern ones.

Traditional pricing strategies originate from insurers knowing their clients personally, and as such the characteristics of their claims. Current strategies mostly rely on established (but old-fashioned) statistical techniques that set the premium on a group level, based on data provided by the policyholders, and are often not regularly recalibrated. Therefore, premiums are not optimized.

The more insight you have in the characteristics of your policyholders, the better you can estimate the corresponding individual risks, and tailor the premium

accordingly. The increasing availability of data facilitates this. We consider the following criteria to be fundamental to fully utilize the advantages:

- a scalable IT infrastructure;
- internal data of high quality;
- the possibility to combine internal data with external datasets (open and commercial);
- updating pricing strategies with modern analysis and modeling techniques, like machine learning.

These points open the door to set a competitive premium on a more individual basis, and updating it directly when new claim data is obtained, or when changes occur in an individual profile. We call this approach a data-driven pricing approach, and strongly believe that this approach will quickly render traditional approaches obsolete. In order to facilitate the change, KPMG has developed an open source analysis platform that meets the criteria set out above, which can be used by insurers to clean, combine and analyze their data. We have described this pricing analysis platform in the box about ‘KPMG open source analysis platform’.

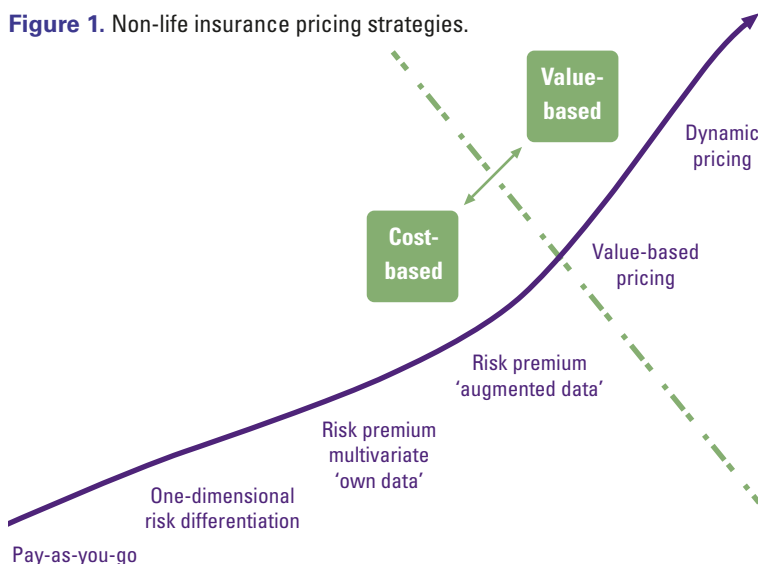
PRICING STRATEGIES

Traditional strategies

Traditionally, insurance companies used to have unique portfolios, for example due to their strong regional presence in the market. People bought insurance from insurance companies present in their area, resulting in portfolios with unique characteristics for the region. Insurance companies used to know the people they insured personally, and could set adequate premiums based on this. With the introduction of buying and comparing insurance policies on the internet, these geographic dependencies started to dissipate, and competition between insurance companies increased. Nowadays it is easy to find ‘the best deal’ for one’s specific insurance needs and buy a policy in a matter of seconds, without any human interaction. For the insurers this results in more geographically dispersed portfolios, with indifferent characteristics. Therefore, these traditional strategies evolved to forms that are supported with data provided by the policyholders and established statistical methods using this data: generalized linear models (GLMs), which were introduced in 1972.

While the market is changing, many insurance companies still apply traditional or old-fashioned GLM-based methodologies to derive their premiums. So how can an insurer set a competitive premium nowadays? For this, we distinguish between cost-based pricing approaches and value-based pricing approaches, which are set out in Figure 1.

Figure 1. Non-life insurance pricing strategies.



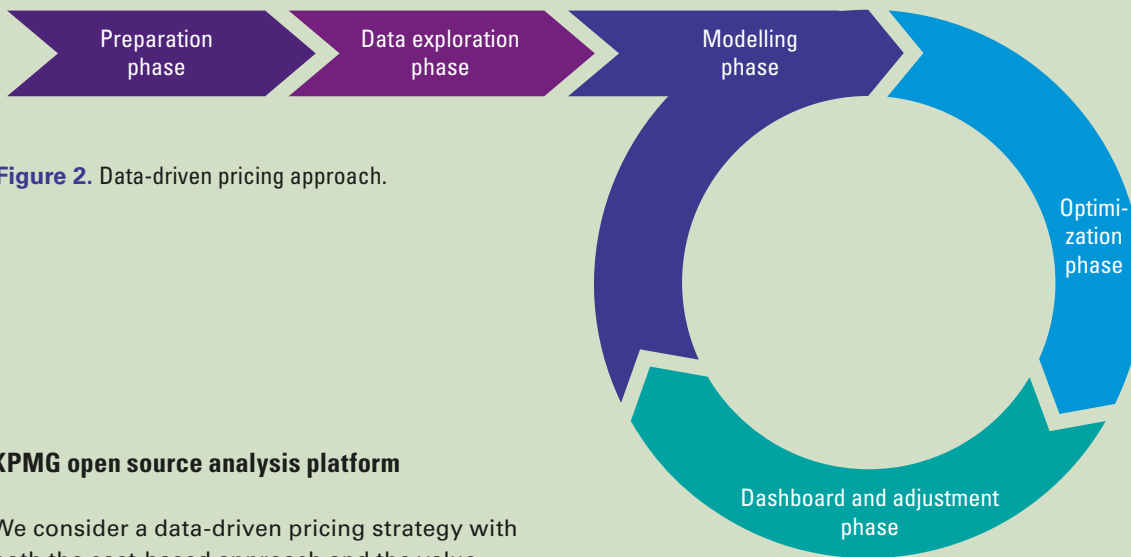


Figure 2. Data-driven pricing approach.

KPMG open source analysis platform

We consider a data-driven pricing strategy with both the cost-based approach and the value-based approach as the new market standard for pricing non-life insurance policies. Viable premiums arise from combining:

1. high quality data from internal and external sources;
2. modern day analysis and modeling techniques;
3. management information.

To support insurance companies in the transition to a data-driven pricing strategy, KPMG has developed an open source-pricing platform: the platform does not, and will not, bear any licensing fees. On this platform modern day data handling and visualization capabilities are combined with up to date modeling techniques to derive an optimal premium. It features a functionality to:

- combine datasets;
- analyze and clean the datasets (data quality checks, pre-production);
- derive premiums with traditional GLM techniques to gain insight on a cost-based approach;
- analyze correlation structure of all variables to enrich the GLMs with cross variables and higher order terms;
- optimize premiums in a value-based approach, e.g. by means of pricing elasticities or competitor information;
- derive premiums for a cost-based and value-based approach with modern machine learning and deep learning techniques;
- visualize portfolio statistics (claims, risk profiles, etc.);
- visualize pricing performance (old/new premiums versus claims).

The platform can easily be deployed by the insurance companies own data scientist, and can be tailored to the organization. The premium can then be optimized following the approach in Figure 2.

- In the preparation phase initial data quality checks are performed and datasets are combined for analysis.
- During the data exploration phase the combined data is analyzed to find well-suited input parameters for the modeling phase.
- Both traditional and modern models are applied to the combined dataset in the modeling phase.
- In the optimization phase parameters are adjusted to find the statistical optimal premium.
- The results are visualized in the dashboard and adjustment phase.
- Based on the statistical and visual interpretation of results and management information, the models can be fine-tuned iteratively.
- New data can be added on a real-time basis, can be used to update the pricing models frequently, and provides direct insight in portfolio developments.

KPMG has experience and credentials in assisting non-life insurers in every step of the way in implementing the approach above, leading to optimized premiums and portfolios. In this matter KPMG offers support from multidisciplinary teams with expertise from both a data and analytics perspective, as well as an actuarial perspective.

Cost-based data-driven pricing strategies

A cost-based pricing strategy involves investigating the risk of a policy, and finding a premium that will cover this risk. This can be done from pooling risk without diversification (pay-as-you-go), to trying to find unique characteristics that forecast the underlying risk of the policy (risk premium augmented data). The latter requires reliable data as input for statistical models. The more unique characteristics can be used, the more individual the risks can be estimated, and therefore the more individual the premium can be set to better cover the risks.

Insurance companies already possess a large amount of data on their policyholders. The digitalization of insurance even enables them to collect more than before, and store the data in a more structured and centralized way. In addition, the availability of open and commercial datasets is ever increasing. For example, the Netherlands Vehicle Authority (RDW) publishes an open dataset with a substantial amount of vehicle information, and the Dutch central agency for statistics publishes regional and demographic information. Using such additional datasets can enhance the insurers' data, by missing data imputations or replacing low quality data in its own set, and enriching it by adding new variables to the set. This provides the insurer extended insights into characteristics of its policyholders and insured objects. However, in our experience insurance companies often face technical difficulties with combining datasets due to different data formats, legacy systems, and an absence of a uniform data analysis platform. If this problem can be overcome and additional data can be used in the pricing procedure, even old-fashioned statistical techniques (like GLM) can be used to set a more precise premium than premiums solely based on internal data. By still finding unique characteristics in the insured portfolio, it is possible to outperform the market with a traditional cost-based pricing approach.

Access to substantial amounts of relevant data also opens the possibility to utilize modern analysis and modeling techniques, like machine learning and deep learning. With these techniques an insurer can extract extensive intelligence from its data in an almost fully automated matter, if desired. Where human interpretation of statistics and modeling choices based thereon are key in traditional methods, these modern techniques automatically find less obvious, but significant causal relationships in the data, which are easily overlooked in a manual assessment. This ultimately results in even more precise and individual premiums. One should remain wary of undesired effects when applying a fully automated approach, like indirect discrimination. For example, regulation prohibits the usage of gender as a pricing

variable for motor insurance policies, but very granular car type and color variables might (partially) act up as a proxy for gender effects.

Value-based data-driven pricing strategies

The methods discussed so far derive the premium with a cost-based approach. They all aim to predict future claims of a policyholder based on certain characteristics, so that an accurate premium to cover these expected claims can be quoted. Another data-driven pricing approach is the concept of value-based pricing. This approach derives the optimal premium based on a customer's willingness to pay for the product. The willingness to pay can roughly be divided in product benefits, service benefits, and brand benefits. Outperforming the competition on these elements can attract new clients, even if the price is higher. Awareness of what the target audience is willing to pay per element is key in optimizing profits. Also, for the value-based pricing approach, the availability of data and modern analysis and modeling techniques are of utmost importance.

CONCLUSION

The pressure and tight margins in the Dutch non-life insurance market forces the market to move to a new pricing approach. Many insurance companies still apply traditional methodologies to derive their premiums. Implementing a data-driven pricing strategy can outperform traditional strategies, and can either be achieved by optimizing the cost-based pricing approach by means of using modern analysis and modeling techniques, but also by further development of the value-based pricing approach.

About the authors

drs. R. Valkenburg MSc AAG is a senior manager at KPMG Financial Risk Management. Rinze is a certified actuary and highly experienced in the insurance industry, mainly on strategic, pricing and reserving topics.

P. Bosschaart MSc is a senior consultant at KPMG Financial Risk Management. Peter is a financial mathematician with five years of experience in the actuarial sector, and has a strong focus on quantitative advisory projects.