BIG DATA ANALYTICS IN MOTOR AND HEALTH INSURANCE:

A THEMATIC REVIEW

https://eiopa.europa.eu/

European Insurance and
Occupational Pensions Authority

Cover photo: istock.com/SusanneB

PDF	ISBN 978-92-9473-142-5	doi:10.2854/54208	EI-02-19-220-EN-N
Print	ISBN 978-92-9473-143-2	doi:10.2854/448071	EI-02-19-220-EN-C

Luxembourg: Publications Office of the European Union, 2019

BIG DATA ANALYTICS IN MOTOR AND HEALTH INSURANCE: A THEMATIC REVIEW

TABLE OF CONTENTS

TABLE OF FIGURES

1. EXECUTIVE SUMMARY

Data processing has historically been at the very core of the business of insurance undertakings, which is rooted strongly in data-led statistical analysis. Data has always been collected and processed to inform underwriting decisions, price policies, settle claims and prevent fraud. There has long been a pursuit of more granular datasets and predictive models, such that the relevance of Big Data Analytics (BDA) for the sector is no surprise.

In view of this, and as a follow-up of the Joint Committee of the European Supervisory Authorities (ESAs) cross-sectorial report on the use of Big Data by financial institutions,[1] the European Insurance and Occupational Pensions Authority (EIOPA) decided to launch a thematic review on the use of BDA specifically by insurance firms. The aim is to gather further empirical evidence on the benefits and risks arising from BDA. To keep the exercise proportionate, the focus was limited to motor and health insurance lines of business. The thematic review was officially launched during the summer of 2018.

A total of 222 insurance undertakings and intermediaries from 28 jurisdictions have participated in the thematic review. The input collected from insurance undertakings represents approximately 60% of the total gross written premiums (GWP) of the motor and health insurance lines of business in the respective national markets, and it includes input from both incumbents and start-ups. In addition, EIOPA has collected input from its Members and Observers, i.e. national competent authorities (NCAs) from the European Economic Area, and from two consumers associations.

The thematic review has revealed a strong trend towards increasingly data-driven business models throughout the insurance value chain in motor and health insurance:

- Traditional data sources such as demographic data or exposure data are increasingly combined (not replaced) with new sources like online media data or telematics data, providing greater granularity and frequency of information about consumer's characteristics, behaviour and lifestyles. This enables the development of increasingly tailored products and services and more accurate risk assessments.
- The use of data outsourced from third-party data vendors and their corresponding algorithms used to calculate credit scores, driving scores, claims scores, etc. is relatively extended and this information can be used in technical models.
- BDA enables the development of new rating factors, leading to smaller risk pools and a larger number of them. Most rating factors have a causal link while others are perceived as being a proxy for other risk factors or wealth / price elasticity of demand.
- BDA tools such as such as artificial intelligence (AI) or machine learning (ML) are already actively used by 31% of firms, and another 24% are at a proof of concept stage. Models based on these tools are often correlational and not causative, and they are primarily used on pricing and underwriting and claims management.
- Cloud computing services, which reportedly represent a key enabler of agility and data analytics, are already used by 33% of insurance firms, with a further 32% saying they will be moving to the cloud over the next 3 years. Data security and consumer protection are key concerns of this outsourcing activity.
- Up take of usage-based insurance products will gradually continue in the following years, influenced by developments such as increasingly connected cars, health wearable devices or the introduction of 5G mobile technology. Ro-bo-advisors and specially chatbots are also gaining momentum within consumer product and service journeys.
- There is no evidence as yet that an increasing granularity of risk assessments is causing exclusion issues for high-risk consumers, although firms expect the impact of BDA to increase in the years to come.

In view of the evidence gathered from the different stakeholders, EIOPA considers that there are many opportuni-ties arising from BDA, both for the insurance industry as well as for consumers. However, and although insurance firms generally already have in place or are developing sound data governance arrangements, there are also risks arising

1 Report on the use of Big Data by financial institutions, Joint Committee of the ESAs, 15 March 2018, https://eiopa.europa.eu/Publications/Other%20Documents/JC-2018-04%20Joint%20Committee%20Final%20Report%20on%20Big%20Data.pdf

from BDA that need to be further addressed in practice. Some of these risks are not new, but their significance is amplified in the context of BDA. This is particularly the case regarding ethical issues with the fairness of the use of BDA, as well as regarding the accuracy, transparency, auditability, and explainability of certain BDA tools such as AI and ML.

Going forward, in 2019 EIOPA's InsurTech Task Force will conduct further work in these two key areas in collaboration with the industry, academia, consumer associations and other relevant stakeholders. The work being developed by the Joint Committee of the ESAs on AI as well as in other international fora will also be taken into account. EIOPA will also explore third-party data vendor issues, including transparency in the use of rating factors in the context of the EU-US insurance dialogue. Furthermore, EIOPA will develop guidelines on the use of cloud computing by insurance firms and will start a new workstream assessing new business models and ecosystems arising from InsurTech. EIOPA will also continue its on-going work in the area of cyber insurance and cyber security risks.

2. TYPES OF DATA AND BDA TOOLS

This section covers the different types of data and data sources used by insurance undertakings and intermediaries in their motor and health insurance lines of business, with a particular focus on the new types of data emerging from digitalisation. It then assesses which types of new data analytics tools (e.g. artificial intelligence and machine learning) are being used to process the datasets, as well as the level of penetration of cloud computing technology in insurance.

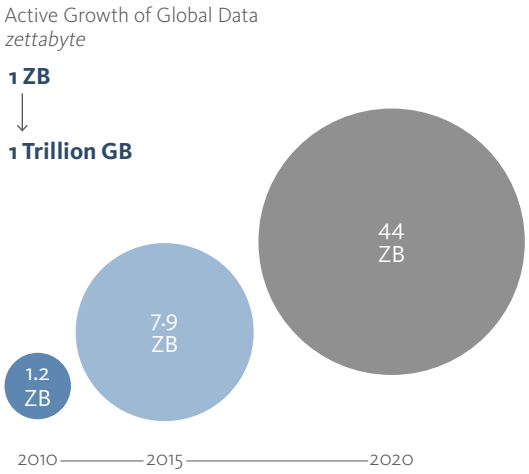
2.1. TYPES OF DATA USED IN INSURANCE

Data processing has always been at the very core of insurance business; traditional datasets such as demographic data, exposure data or behavioural data have historically been processed by insurance firms to inform underwriting decisions, price policies, evaluate and settle policyholders' claims and benefits, as well as to detect and prevent fraud. In the era of digitalisation, these traditional datasets are increasingly combined with new types of data such as Internet of Things (IoT) data, online data, or bank account / credit card data in order to perform more sophisticated and comprehensive analysis, in a process that is commonly known as 'data enrichment.'

The data used by insurance firms in the different stages of the insurance value chain may include personal data[3] (e.g. medical history) as well as non-personal data (e.g. hazard data), and it can be structured (e.g. survey, IoT data) or unstructured (e.g. pictures or e-mails). It can be obtained from internal sources (e.g. provided directly by the consumer to the firm) as well as from external sources (e.g. public databases or private data vendors).

The table below aims to cluster into different categories the types of data used by insurance firms, which is followed by a more detailed assessment of the uses and sources of these types of data and the extent to which they are already used or not in the insurance sector.

Figure 1 – Increasing availability of data



Source: Institute of International Finance[2]

2 Report "Innovation in Technology: How technology is changing the industry", Institute of International Finance, September 2016, https://www.iif.com/system/files/32370132_insurance_innovation_report_2016.pdf

3 According to Article 4 GDPR, personal data is any information relating to an identified or identifiable natural person; an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.

Figure 2 – Traditional data sources and new data sources enabled by digitalisation

Traditional data sources	New data sources enabled by digitalisation
Medical data (e.g. medical history, medical condition, condition of family members)	**IoT data** (e.g. driving behaviour (car telematics), physical activity and medical condition (wearables).
Demographic data (e.g. age, gender, civil and family status, profession, address)	**Online media data** (e.g. web searches, online purchases, social media activities, job career information)
Exposure data (e.g. type of car, value of contents inside the car)	**Insurance firms' own digital data** (e.g. interaction with insurance firms (call centre data, users' digital account information, digital claim reports, online behaviour while logging in to insurance firms' websites or using insurance firms' app)
Behavioural data (except IoT data) (e.g. Smoking, drinking behaviour, distance driven in a year)	**Geocoding data** (i.e. latitude and longitude coordinates of a physical address)
Loss data (e.g. claim reports from car accidents, liability cases)	**Genetics data** (e.g. results of predictive analysis of a person's genes and chromosomes)
Population data (e.g. mortality rates, morbidity rates, car accidents)	**Bank account / credit card data** (e.g. consumer's shopping habits, income and wealth data)
Hazard data (e.g. frequency and severity of natural hazards)	**Other digital data** (e.g. selfie to estimate biological age of the consumer)
Other traditional data (e.g. credit scoring, claim adjustment reports, information from the auto repair shops)	

Source: The Geneva Association (the categorisation of types of data was slightly amended by EIOPA)[4]

2.1.1. TRADITIONAL DATA SOURCES

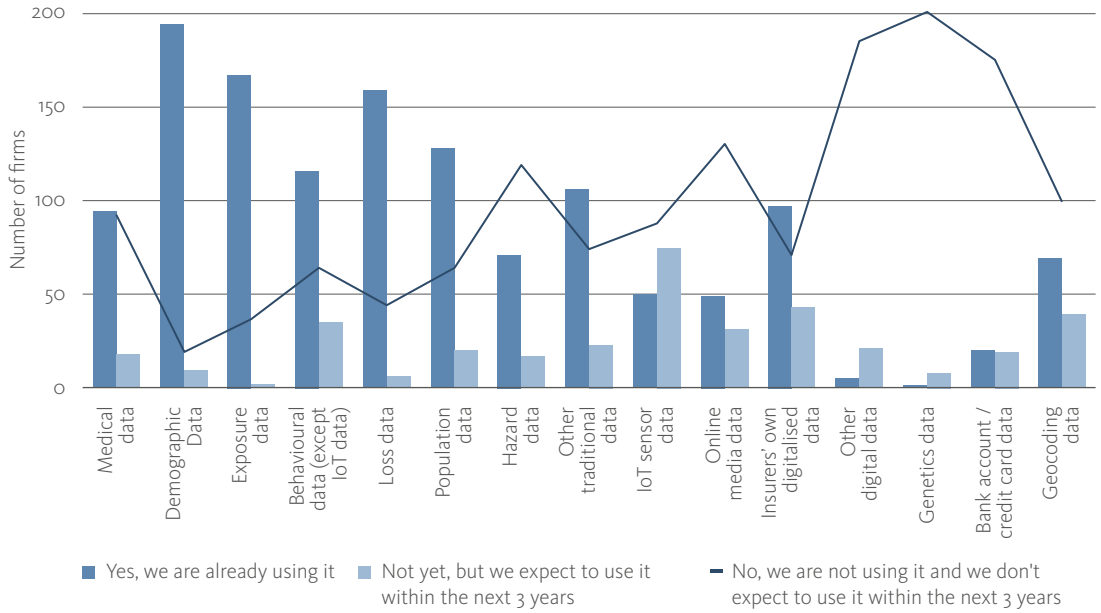
For health insurance purposes, medical data is commonly collected directly from the customer or via insurance intermediaries. This is often done by completing a medical survey or undergoing an examination by a medical institution to assess the customer's medical condition. Some insurance firms also use statistics about costs of disease published annually by the relevant NCA or the national association of private health insurance firms.

Some of the respondents to EIOPA's survey specifically stated that they do not use medical data in their BDA processes. Others referred to the outsourcing to service providers the use of powerful digital tools capable of processing a significantly greater volume of health information than traditional underwriting tools. Moreover, one firm planned exporting the medical history of their customers into their health claims app in order to improve their customer's user experience.

Finally, motor insurance firms also collect medical data to assess bodily injuries arising from car accidents. At the pre-contractual stage, some insurance firms also ask consumers to provide information about any special medical conditions that could affect their driving capacity.

4 In order to categorise the types of data used by insurance firms, EIOPA has used as a basis (and slightly amended) the classification included in the Appendix of the Geneva's Association report titled: "Big Data and Insurance: Implications for Innovation, Competition and Privacy", March 2018, https://www.genevaassociation.org/research-topics/cyber-and-innovation-digitalization/big-data-and-insurance-implications-innovation

Figure 3 – Usage of different types of data



Source: EIOPA BDA thematic review

Almost all insurance firms answering the questionnaire use demographic data (e.g. age, gender, occupation, civil status etc.) both from the motor and health insurance sectors, which is commonly obtained directly from the consumer at the pre-contractual stage. The information provided by the customer is often complemented with external data sources such as national statistics offices or third-party data vendors.

For instance, some firms explained how they receive geo-spatial socioeconomic demographic classifications (e.g. purchasing power, family types, population density etc.) from third-party data vendors. This information is often provided at postal code level (i.e. anonymised) and increasingly at a more granular level (see below the use of geocoding data) and it is commonly used in technical models for pricing and underwriting purposes. Insurance firms also collect gender data. However, they are not allowed to use it for pricing and underwriting purposes following the 2011 ruling of the European Court of Justice against the pricing differentiation on the grounds of sex.[5]

Exposure data is the second most common category of data used by insurance firms, fundamentally from the motor insurance sector; consumers typically have to provide information about their car when entering into an insurance contract. More detailed information about the vehicle (e.g. brake horsepower, acceleration, height, weight) and its current value can be obtained from private external sources using the car's license plate and/or its registration number. In some jurisdictions, these types of non-personal information can be obtained from public sources such as national traffic authorities or the motor insurance bureau.

The majority of motor and health insurance firms also use behavioural data (other than IoT data). In most cases, this data is obtained during the onboarding of the customer, although it can also be obtained from external sources such as criminal statistics provided by local or national governments. Behavioural data can also be obtained after the onboarding of the customer, for instances regarding the delays in instalments. Moreover, some respondents mentioned that in their jurisdictions, the vehicle is the insured object (i.e. not the driver) and therefore in their opinion data such as number of kilometres driven in a year is not considered as personal data because the 'behaviour' is of the car.

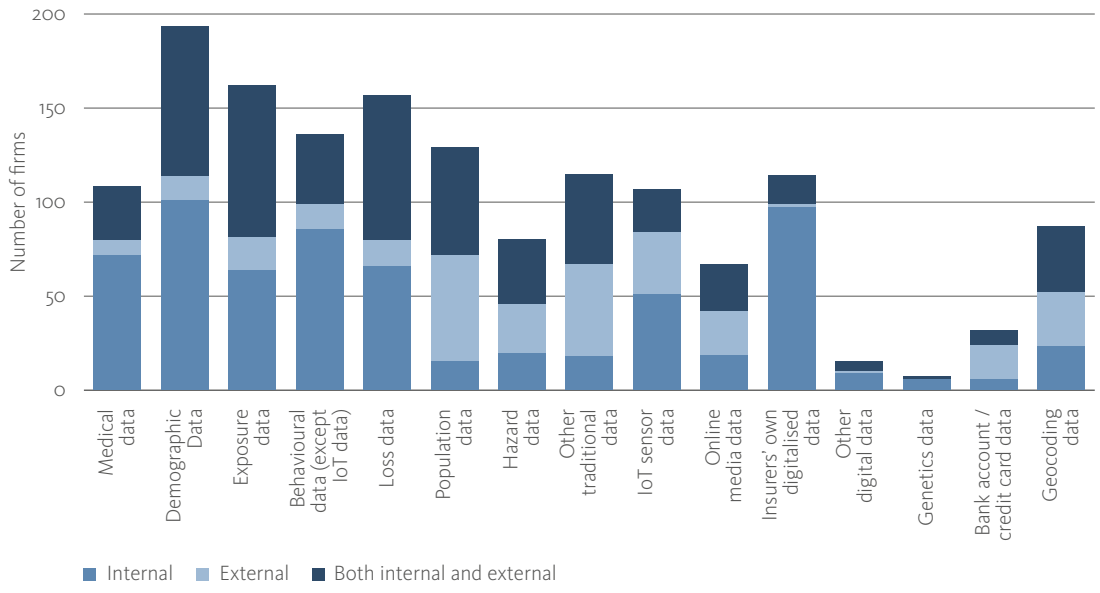
5 Case C-236/09, The European Court of Justice, 1 March 2011, http://curia.europa.eu/juris/document/document.jsf;jsessionid=9ea7d0f130d0a5e2021e6604e42a81bb9c9e243b89a.e34KaxiLc3eQc40Lax-qMbN4PaNyPe0?text=&docid=80019&pageIndex=0&doclang=en&mode=lst&dir=&occ=first&part=1&cid=69679

As far as loss data is concerned, this type of data is used both by motor and health insurance firms and it is evenly collected from internal and external sources. When entering into a motor insurance contract the consumer is typically asked to provide its claims history (i.e. Bonus Malus). Loss data is also captured for claims settlement purposes during the first notice of loss (FNOL) by completing the relevant claims forms. It is a common practice by insurance firms to validate the data provided by the consumer via public sources (e.g. motor insurance bureau), external data vendors or incidentally via loss adjusters or private investigators which may assess different sources such as photographic evidence, video surveillance (CCTV), police reports, witnesses testimonies etc. Some firms reportedly have outsourced the claims management phase to specialised third parties. Health insurance firms reported the use of information such as historical absenteeism of employees of a given firm or medical treatment statistics in different medical institutions (e.g. non-personal information about costs and treatments).

Population data is mainly collected in an aggregated manner (i.e. not personal information) from public sources such as national statistics office or the Ministry of Internal Affairs or the insurance bureau, the national actuarial association or even the World Health Organisation. Some firms outsource the analysis of this publicly available information to service providers which typically provide trends analysis at post-code level; population data, including census data, can be used to inform post-code classifications.

Similarly, hazard data is predominantly obtained from public sources such as the national weather institute, although it can also be sourced from the consumer, namely during the FNOL. The use of this data is relatively small compared to other types of data, especially amongst health insurance firms. Insurance firms often outsource natural catastrophe models to reinsurance brokers and other service providers.

Figure 4 – Internal and external data sources



Source: EIOPA BDA thematic review

Regarding 'other traditional data', several firms declared using credit scoring reports both for claims management (including fraud prevention) as well as for pricing and underwriting. These reports are predominantly obtained from specialised third-party data vendors, although some firms also obtain this type of information from their bank assurance distribution channel. This information is commonly assessed in combination with other internal types of data such as the internal claims records and premium payment history for existing customers available at the firm or also in occasions provided by the motor insurance bureau.

Finally, the other main type of 'other traditional sources' is the information provided by auto-repair shops and claims adjustment reports. Firms use these reports for a variety of purposes, such as assessing trends in repair costs or in order to improve vehicle classifications (e.g. average cost of claims per car brand, model etc.).

2.1.2. NEW TYPES OF DATA ENABLED BY DIGITALISATION

Of the 222 insurance firms that participated in EIOPA's survey, 50 firms already use IoT data, while another 75 of them expect to use it within the next three years. The use of this type of data is more extended in the motor insurance sector, either via black boxes installed in cars or mobile phone apps. Reported examples of data collected through these devices include geolocation data, speeding data, miles driven, harsh braking, time of day, road type, g-forces etc. This information can be complemented with external data such as the speed limit and type of street in order to assess whether the consumer respects the driving rules.

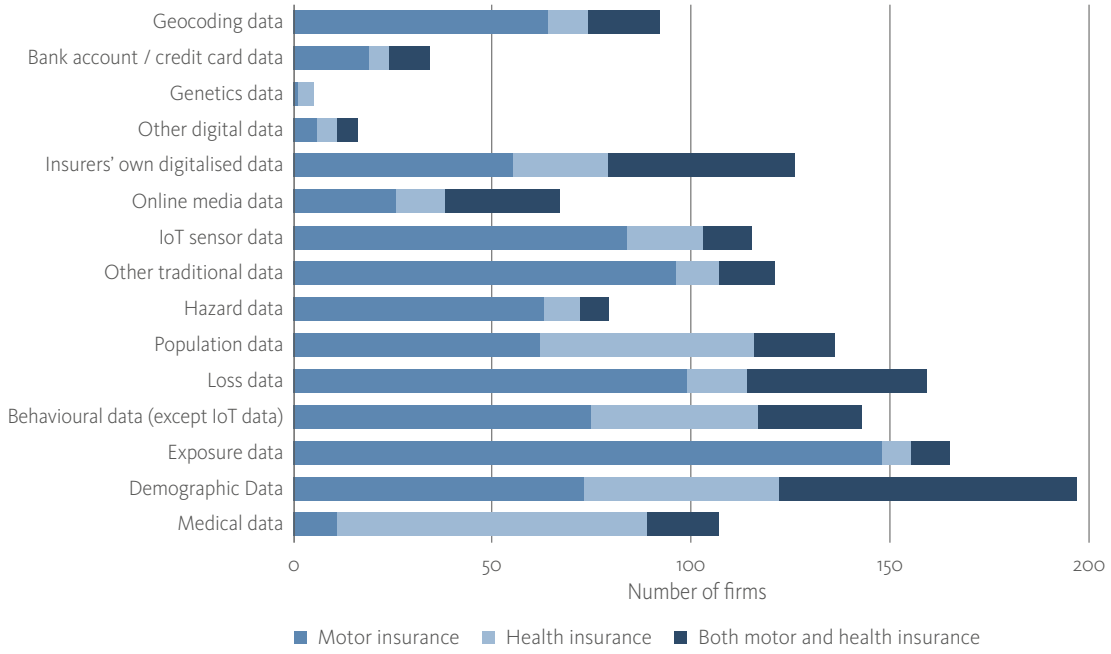
A reduced number of health insurance firms also use wearable devices and mobile phone apps to collect consumer information such as life activity, steps walked in a day, calorie consumption, blood pressure etc. The information collected from these devices is obtained both from internal and external data sources and is mainly used for product development and sales and distribution, although some of them also use it for pricing and underwriting and claims management.

Several insurance IoT projects are currently on a testing phase; for example, one firm is testing IoT health insurance products with its employees. Moreover, telematics offers often target young customers; young drivers installing a telematics device in their car reportedly have access to more affordable motor insurance. When used for pricing and underwriting (e.g. Pay-As-You-Drive or Pay-How-You-Drive policies), the consumer typically receives a driving score or a health score derived from the information gathered through the telematics device. Telematics devices can also be used to offer customer emergency call systems; in case of an accident, the system collects data about the severity of the brunt and its geolocation and automatically sends an emergency call and forwards this information to the firm's claims department.

Insurance firms use their own digitalised data for several purposes, such as analysing the consumer's behaviour during the quote process (e.g. contact centre logs and website cookies) for fraud prevention purposes and for optimising customer journeys as well as for pricing/risk modelling, even on a real-time basis. Examples of information assessed include trends in distribution channels used, number of contacts, how much time the consumer spent reading the terms and conditions, if he/she read it before registering a claim, what directories the consumer uses before buying a policy or making a claim etc.

This information can also be used to trigger marketing automated practices to target specific customers if certain behavioural patterns arise in their website or app. Moreover, information relevant for claims management purposes (e.g. medical bills images or images of damaged cars) are also increasingly submitted via digital means.

Figure 5 – Usage of data in the different lines of business



Source: EIOPA BDA thematic review

Traditional postal codes (which could be a proxy of wealth,[6] crime, population density or ethnicity) used for pricing and underwriting are being substituted/combined by more accurate 'micro-zoning' data sourced both from external providers as well as from apps and other telematics devices. For example, one firm uses geocoding data provided by a third-party data vendor to append on socio demographic area-related information. Firms can also use this information to analyse areas with increased claims expenses (e.g. hail damage or car theft). Indeed firms explained how the consumer's place of residence or work influences applicable risks.

Moreover, the latitude and longitude coordinates of a physical address are also used to measure the distance from other points of interest such as schools, shopping centres, commute stations, hospitals, major transport hubs, accident hotspots etc. Some insurance firms use this accurate geocoding data to optimize the network of repair shops and distribution channels. Some firms also use the geocode ob-tained from the tow truck driver arriving at the site of the car accident.

Bank account and credit card data is already used by 20 insurance firms from nine countries, while 19 more from nine other countries said that they expect to use it within the next three years. This information is gathered both from internal and external sources; for example one firm explained how they use customer segmentation based on shopping information gathered from retail store's loyalty cards. Another firm described how they use bank account and credit card information to predict churn.

Another firm said that they expect to collect this information via a third-party leveraging on the greater possibilities to access banking data offered by the recently approved

6 For example, since December 2018 the Spanish Tax Authority ('Agencia Tributaria') publishes the average wealth of all of the municipalities in Spain based on their postal code; https://www.agenciatributaria.es/AEAT.internet/datosabiertos/catalogo/hacienda/Estadistica_de_los_declarantes_del_IRPF_por_municipios.shtml.

Payments Services Directive (commonly known as 'PSD 2').[7] Another firm uses an external provider to prevent fraud by checking if a credit card has been used with multiple different names. It is also possible to create different scorecards based on bank information, shopping habits and/or mobile phone use. Some bank-assurance business models use this information to provide a discount in motor insurance based on their own credit scoring mechanisms. Moreover, one firm explained how they use the external data provided by entities such as credit rating agencies or shopping stores to construct accurate socio-demographic profiles at postcode, household and individual level.

In addition to the digital data from their own websites, insurance firms can use other third-party online media data. Indeed aggregated information like search engine searches and website visit information[8] is frequently used by the marketing departments to customize campaigns. Information from online comparison platforms can also be used to improve sales and distribution practices. Some insurance firms also track (public) social media posts for counter-fraud services, provided either by specialist counter-fraud service or by specialised employees not necessarily with the use of BDA tools.[9] Some firms also reportedly use the information provided by web analytics tools and social media listening tools for analysis of user behaviour and for pricing and underwriting purposes.

Finally, only one health insurance firm declared that they already use genetics data, although nine firms from seven Member States expect to use it within the next three years.

2.2. BIG DATA ANALYTICS TOOLS

2.2.1. ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

It would not be possible to understand the phenomenon of Big Data without the use of increasingly powerful IT tools, algorithms and information systems to make more predictive, descriptive and prescriptive analysis. This is the case of artificial intelligence (AI) tools, which among others can autonomously discover and test hypotheses and make decisions automatically. AI algorithms can also access previously inaccessible datasets such as unstructured data from pictures, videos or audios.

Machine learning (ML) is one subcategory of AI, which refers to sophisticated computer algorithms that have the ability to find optimal solutions, across increasingly large data sets, without a traditional rules-based approach. For the insurance industry, these techniques coupled with increasing amounts of data (for example through increased interconnectivity from telematics and wearable devices) could represent a step-change improvement in the ability to accurately monitor and evaluate risk.

MACHINE LEARNING: ARTIFICIAL NEURAL NETWORKS

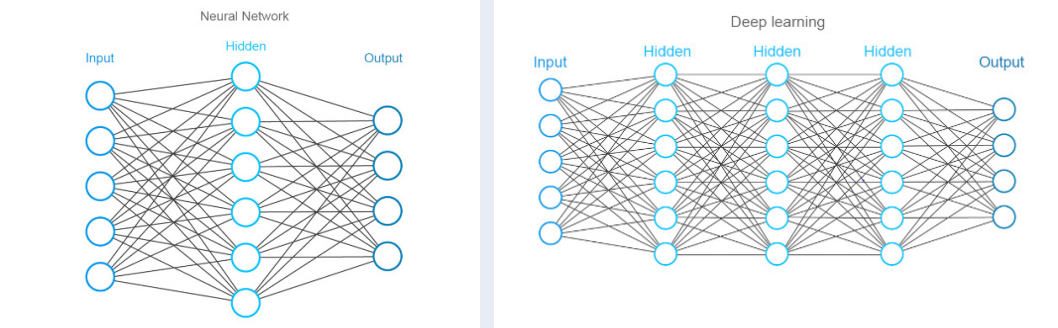
One currently widespread type of machine learning algorithm, which has already existed for a number of years, is artificial neural networks (ANNs). ANNs are algorithms that are commonly represented as working in a similar fashion to the human brain; they operate with an input layer (e.g. claims history, age of the driver, address, etc.), one or more unknown hidden layers, and an output layer (e.g. consumer does or does not renew the contract etc.). In what is known as a 'feedforward network', the information flows from the input layer, through the hidden layer(s) into the output layer.

7 Directive (EU) 2015/2366 of the European Parliament and of the Council of 25 November 2015 on payment services in the internal market, amending Directives 2002/65/EC, 2009/110/EC and 2013/36/EU and Regulation (EU) No 1093/2010, and repealing Directive 2007/64/EC, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32015L2366

8 Like the one provided by Google Analytics and SEO tools

9 The Danish NCA conducted a thematic review on this topic in 2016 and published a report describing the applicable legal framework. The report specifies that using false profiles on social media or using GPS trackers to monitor the activity of individuals are practices that are not considered to be in line with the insurance firms' obligations to treat their clients fairly. http://www.finanstilsynet.dk/da/Nyheder%2009%20presse/Pressemeddelelser/2016/Pressemedddelelse-rapport-om-forsikringsselskabers-efterforskning-ved-personskade-030216

Figure 6 – Artificial Neural Networks and Deep Learning Networks

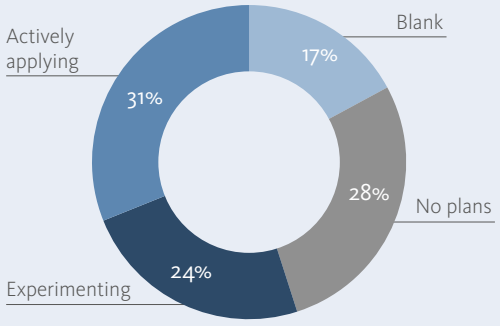


Source: Altexsoft*

ANNs can be 'trained' in order to adjust the individual weights of the hidden layers to the desired output. This learning process commonly includes the method of 'backpropagation,' where the output layers are compared with the correct answers in order to input error values into the network until the errors in prediction are low. ANNs can run this process very efficiently finding decisions that are not subject to specific structural conditions, especially when trained with large quantities of data.

The penetration of BDA tools such as ANNs is still limited in the insurance sector; although 55% of the firms claim to have already started using them or plan to do it within the next three years; only 30% of them are already actively applying BDA tools, as shown in Figure 7 below.

Figure 7 – Usage of BDA tools such as AI and ML



Source: EIOPA BDA thematic review

The main benefits of ANNs and ML more broadly are that the algorithms are able, in certain domains at least, to develop, without or with limited human intervention, very accurate assessments, increasing the efficiency and speed of decision making and therefore reducing operational costs. However, this also causes specific challenges in terms of accuracy, transparency, explainability and auditability of the models using ML algorithms, which are often correlational and not causative; due to the opacity of the functioning of the 'hidden layer', it can be difficult to explain the (causal) relationship between the inputs and outputs of the model. The higher the complexity of the model (e.g. see deep learning models in Figure 6), the greater these challenges are.

In particular, fair use of ML algorithms relies on historic data that is without demographic slant, since otherwise any biases inherent in the historic data will be reinforced through the ML algorithm if there are no adequate governance arrangements to address these issues. It should be noted that 'traditional' generalised linear models can become so complex that they arguably are in practice not easier to explain than a more advanced ML model. Therefore, sound governance arrangements should always be in place irrespective of the type of model used.

* Fraud Detection: How Machine Learning Systems Help Reveal Scams in Fintech, Healthcare, and eCommerce, Altexsoft, https://www.altexsoft.com/whitepapers/fraud-detection-how-machine-learning-systems-help-reveal-scams-in-fintech-healthcare-and-ecommerce/

According to the responses to the questionnaire received by EIOPA, BDA tools are more commonly used by insurance undertakings than by insurance intermediaries.[10] BDA tools are equally used in motor and health insurance lines of business.[11] Additionally, most of the firms that use BDA have developed the solution in-house, although many others have bought them 'off-the-shelf' from third-party service providers. Several firms also mentioned that they use or are starting to use open source (i.e. freely available) tools based programming languages such as 'Python' or 'R' for these purposes.

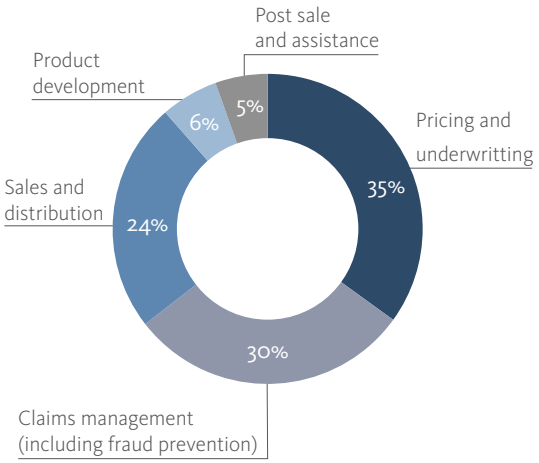
The use of BDA tools is generally focussed on a specific part of the insurance value chain and very few firms make use of them across all their processes. Amongst those firms that already use or are planning to use BDA tools, they mainly use them for pricing and underwriting, claims handling and sales and distribution, as it can be observed in Figure 8.

Figure 8 shows that to date BDA has had a greater impact in the pricing and underwriting area of the insurance value chain, which is also in line with the expected evolution of these technologies within the next three years.

10 20% insurance intermediaries and 59% of insurance undertakings reported using or planning to use BDA tools

11 59% of insurance firms active only in health insurance line of business and 59% of insurance firms only active in motor insurance line of business reported using or planning to use BDA tools

Figure 8 – Usage of BDA tools such as ML and AI across the value chain



Source: EIOPA BDA thematic review

Figure 9 provides examples of different types of BDA use cases which are already being used or planned to be used by some insurance firms across the insurance value chain (please note that this information is further developed in the next sections of the report).

Figure 9 – BDA uses cases

Use Case	Output
Churn models	Use of ML churn models for the prediction of consumer's propensity to shop around at the renewal stage, which can be useful for pricing and underwriting (e.g. for price optimisation in combinaiton with a demand price-elasticity analysis) or for servicing the customer (e.g. "Next Best Action" approach)
Chatbot	Enable "human like" conversations with consumers by analysing customer unstructured data via text or voice with the use of natural language processing and other ML algorithms
Sentiment Analysis	Evaluate the sentiment in feedback provided by consumers to transform it into usable information to help improve customer satisfaction and engagement
Electronic document management	Robotic process automation (RPA) – Deep learning networks used for automatic classification of incoming documents of unstructured data (e.g. emails, claims statements), routing them to the correct department
Claims management	Optical character recognition (OCR) - Deep learning networks used to extract information from scanned documents such as images from damaged cars to estimate repair costs
Fraud prevention	Analysis of fraudulent claim patterns based on FNOL data provided by the consumer
Product development	Use of ML and graph database in predictive modeling for the identificaiton of disease development patterns
Pricing and underwriting	BDA tools used in motor and health insurance for processing large quantities of data from different sources, often on a real-time basis (e.g. quote manipulation), using a wide array of statistical techniques

Source: EIOPA BDA thematic review

2.2.2. CLOUD COMPUTING

Most financial firms, including the ones from the insurance sector, operate their own corporate data centres that were designed as fit-for-purpose for their specific business needs and BDA processes. In parallel, the use of cloud computing technology has become increasingly widespread since the late 2000's. The reportedly faster time to market, lower development costs, expanded testing, higher apparent resilience, and automatic scaling are some of the features that make cloud computing technology particularly relevant in a BDA context.[12]

Indeed cloud computing technology is an increasingly popular tool enabling the implementation of BDA solutions such as the ones described in the previous point. Moreover, cloud computing service providers (fundamentally Big Tech firms) reportedly also offer to their customers the possibility to outsource AI and ML tools through commoditised computer platforms in parallel to their data storages services.

In the insurance sector, at least 74 firms (i.e. 33% of the total firms) already use at least one of the three main types of services offered by the providers of cloud computing technology, with Software as a Service (SaaS) having the highest adoption rate (28%), followed by Infrastructure as a Service (IaaS) (21%) and Platform as a Service (PaaS) (20%). Many firms claim to use several of these services at the same time for different purposes. Another 72 firms (i.e. 32%) which are not currently using any type of cloud computing service say that they will be moving to the cloud over the next 3 years. Figure 10 shows the current and expected level of adoption of this technology within the next 3 years.

12 For further information on the use of cloud computing by insurance firms, please see 'Outsourcing to the cloud: EIOPA's contribution to the European Commission Fintech Action plan' of March 2019

Figure 10 – Penetration of cloud computing services

Cloud Computing Service	Yes, we are already using it	Not yet, but we expect to use it within the next 3 years	No, we are not using it and we don't expect to use it within the next 3 years	Blank
IaaS	21.2%	21.6%	41.4%	15.8%
PaaS	20.3%	23.0%	41.0%	15.8%
SaaS	28.4%	20.3%	36.5%	14.9%

Source: EIOPA BDA thematic review

When questioned about the obstacles that firms might have found for outsourcing cloud-computing services, several firms raised data security concerns regarding the information stored on the cloud because of personal data protection regulations. Non-anonymized data cannot be processed, it was argued, in the cloud without addressing major obstacles.

These issues seems to be of even higher importance when the data is stored and processed in a third-country outside the EU or if there is a lack of transparency considering the geographical location of the data where cloud computing providers balance data across multiple centres. Some insurance firms also pointed out the refusal of some cloud-computing providers to take into account and include in data processing agreements some specific protection requirements adapted to insurance specificities. Some insurance firms also raised possible vendor lock-in concerns.

Moreover, some insurance firms also consider that regulatory barriers such as legal outsourcing provisions are also hard to overcome. Some firms believe that there is still a high level of uncertainty in areas such as how to handle audit rights, risk management, etc. and therefore consider that a greater clarity in this area would help leveling the playing field between regulated and non-regulated entities.

3. USE OF BDA THROUGHOUT THE INSURANCE VALUE CHAIN (EXCEPT PRICING AND UNDERWRITING)

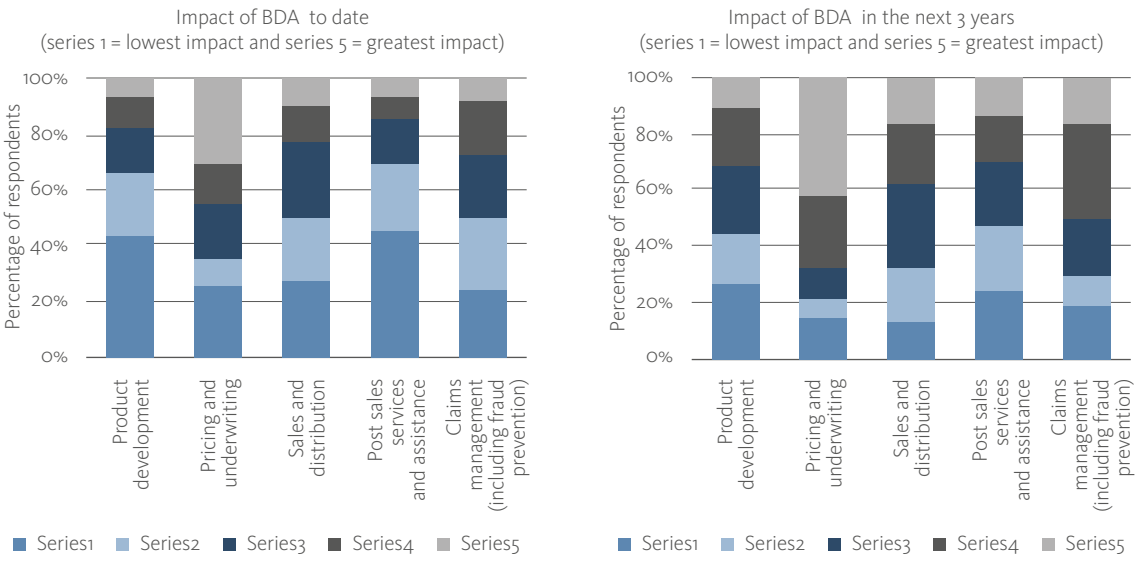
Out of the 222 insurance undertakings and intermediaries that participated in EIOPA's thematic review, the majority of them consider that to date, BDA has had a biggest impact in the pricing and underwriting stage of the insurance value chain, followed by claims management and sales and distribution. The same pattern is reflected when firms are questioned about the expected evolution of BDA within the next three years. Firm's responses also show that they expect data-driven business models to become increasingly significant in all of the areas of the insurance value chain over time (see Figure 11).

3.1. PRODUCT DEVELOPMENT

According to the information available at EIOPA, most insurance firms consider that BDA will enable them to better understand their customer's needs and characteristics and therefore allow them to develop more personalised products and services. Firms expect BDA will transform product development processes and product customization through the ability to identify underlying patterns in extremely granular data, coupled by the ability to capture and use increasingly available behavioral data from consumers.

Insurance firms consider that BDA enables them to develop more granular risk assessments and better segmentation of consumers by means of assessing the risks in areas and segments that was not possible in the past. This results in the definition of new risk factors that enable the development of new products, both in motor and in health insurance, focusing on specific targets, markets and groups of coverage. It was also explained that BDA would allow firms to test the feasibility of new product concepts prior to releasing them to market.

Figure 11 – Impact of BDA across the insurance value chain*



Source: EIOPA BDA thematic review
* Blank responses have not been considered in these graphics

USAGE-BASED INSURANCE PRODUCTS

The positive impact of BDA in the product governance process is particularly relevant in the context of the Internet of Things (IoT). In insurance, this has materialised in the emergence of Usage-Based Insurance (UBI) products in motor and health insurance, i.e. insurance products measuring consumer's behaviour and environment to perform risk assessments and price discount rewards.

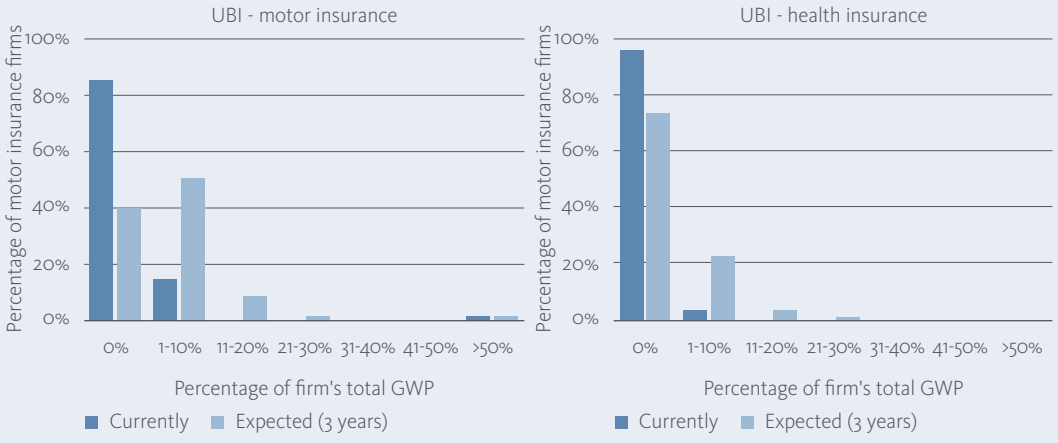
There are two main types of UBI products in motor insurance; on the one hand with the so-called Pay-As-You-Drive (PAYD) policies, the premiums are based on the number of kilometers driven by the consumer. In Pay-How-You Drive (PHYD) policies, consumers receive a driving score depending on their driving behavior (e.g. number of kilometres driven, the average speed, acceleration, geolocation etc.) which influences the final premium paid by the consumer.

Pay-As-You-Live (PAYL) policies in health insurance use wearable devices tracking variations in blood pressure, glucose levels, number of steps walked, calories consumption, places visited etc. which can also be used to perform risk assessments and price health insurance products. Health / lifestyle scores can also be developed. Under these types of policies, consumers demonstrating that they follow healthy lifestyles (e.g. low calorie consumption) receive premium discounts and other types of rewards.

Based on the information gathered by EIOPA from the insurance industry, the current level of penetration of UBI in Europe is still low, especially in the health insurance sector; from the 222 insurance firms that participated in the thematic review, only 15% of the European motor insurance firms and 4% of the health insurance firms currently offer some kind of UBI product (<10% of their total GWP). In the next three years, pos-sibly taking into account the increasingly connected vehicles and the upcoming introduction of 5G mobile technology, 50% of the motor insurance firms and 23% of health insurance firms expect these products to represent up to 10% of their total GWP.

UBI products offer various benefits to both insurance firms and consumers. For firms, this means better-tuned predictive actuarial models leading to reduced claims costs, improved decision-making and increased customer satisfaction. Consumers also benefit from better control of their premium as well

Figure 12 - Penetration of UBI products in motor and health insurance[13]



Source: EIOPA BDA thematic review

13 Blank responses have been interpreted as "0%" in this graphic

as continuous feedback that could help them improve their driving and/or health style behavior. On the other hand UBI also raises a number of challenges, namely from a data-privacy perspective as well as data quality issues which can affect the outputs of BDA tools (e.g. driving data collected from black boxes is reportedly more reliable than the one collected from mobile phone apps).

While the use of telematics is commonly associated with usage-based insurance, it is important to highlight that telematics are not only used for pricing insurance policies. Indeed firms have reported a wide range of services that they offered to consumers through their respective telematics devices, which can broadly be classified between assistance services and risk prevention and mitigation services. These services are not new. For example, health insurance firms often offer to their customers a wide array of diseases management services (e.g. annual medical check-ups). However, while these services were traditionally based on a static set of information and medical criteria, with BDA it is possible to process more granular information on a continuous basis and therefore offer more tailored and timely services to consumers.

Figure 13 shows the different types of telematics devices and data collected in motor and health insurance. An overview of the services is also provided (see also figure 16 about mobile phone apps).

Figure 13 – Types of telematics devices

Line of business	Type of telematics device	Type of data collected (depends on the telematics device)	Types of services offered (depends on the telematics device)
Motor insurance	On board device (OBD) dongle or "black box", mobile phone app, GPS, emergency message plug, forward facing cameras ("dash cams")	Average speed, maximum speed, acceleration and braking habits (G-forces), geolocation, distance travelled, time of travel (e.g. day or night), number of journeys, crash reports, battery and engine condition, cornering, lane changes	**Risk mitigation and prevention**: premium discounts based on driving habits, preventive push-notifications or alerts (e.g. black-spot roads or bad weather conditions or battery and engine breakdown problems), travel statistics reports, driving coach recommendations, treats and vouchers for good driving behaviour **Assistance:** road assistance in case of accident or car theft, emergency call in case of accident (ecall)
Health insurance	Wearable bracelets and other fitness trackers, mobile phone app, smart watch	Heart beat rate, blood pressure, blood oxygen level, activity data (e.g. sports or step counter), hours of sleep, geolocation, food and water consumption, calorie consumption, glucose level.	**Risk mitigation and prevention:** rewards for healthy habits, health activity reports, diabetes management **Assistance**: medical assistance services in case of accident, safety alarm for elderly (e.g. BDA tools can predict falls from anomalies in usage/activity patterns)

Source: EIOPA BDA thematic review

3.2. SALES AND DISTRIBUTION

BDA is widely used in this part of the value chain according to the insurance firms that participated in EIOPA's thematic review. Firms can more precisely profile and segment consumers (e.g. based on marital status, number of insured persons, geolocation, premium size, etc.) enabling them to model cross-selling, up-selling and churn propensities with the help of BDA tools such as AI and ML. The scores resulting from these models are then used to develop increasingly targeted and personalised marketing campaigns and steer sales agent's activities.

Some firms have developed sophisticated customer relationship management (CRM) systems where all the information available about their customers (contact details, number of policies, emails, customer interactions etc.) is integrated into a single platform to support sales management. This allows firms to accurately develop forecasts of business volumes, consumer lifetime value estimations, and define customer loyalty and retention campaigns, sometimes with real-time response capabilities. CRM systems also allows firms to more accurately measure the effectiveness of the campaigns after it has been finalised.

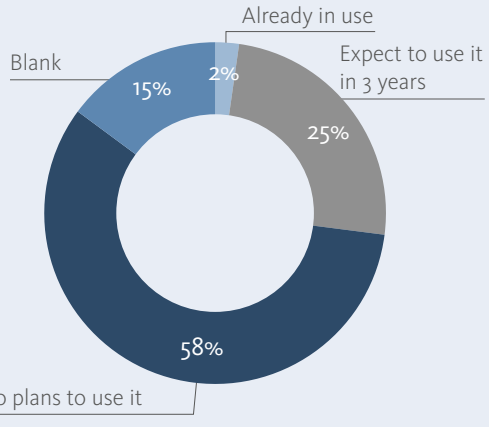
Several firms explicitly referred to the concept of 'Next Best Action', where BDA tools such as ML are used to evaluate the consumer's past behaviour, recent actions and needs in order to deliver the right message, at the right time, and via the right channel. This has been compared to a 'consumers who bought this, might also buy this' approach, allowing the enhancement of consumer's interaction process, fine tuning of advertisements and the determination of the optimal distribution channel mix.

It was also mentioned the potential use of BDA to replace traditional 'form-filling' insurance application processes by the use of third-party service providers, supplemented by images, Customer User Interfaces (CUI) and speech to text capabilities. Some health insurance firms mentioned their intention to use 'look-a-like' marketing modelling where BDA tools are used to analyse the target audience, identify their key characteristics and find other consumers who are similar to the target market. Finally, the use of tools such as Google Analytics and social analytics to track their website activity is also extended amongst firms.

ROBO-ADVISORS

One relevant development in this area of the value chain is the development of robo-advisors, i.e. advice* is provided to consumers without, or with little, human intervention and providers rely instead on computer-based algorithms and/or decision trees. As it can be observed in Figure 14, the level of adoption of this technology is still very low, with only 2% of firms that participated in EIOPA's survey already using them. On the other hand, up to 25% of the firms expect to use them within the next 3 years.

Figure 14 – Robo-advisors penetration in insurance



Source: EIOPA BDA thematic review

The level of adoption of robo-advisors is lower than the one of chatbots (see figure 15). Moreover, amongst those motor and health insurance firms that already use robo-advisors or plan to use them within the next 3 years, most of them have decided to deploy it in cooperation with external service providers. More concretely, 50% have opted or will opt for outsourcing building the solution to a third-party provider, 39% for buying it off-the-shelf from a third-party provider, and only 11% for building it in-house.

The Joint Committee of the ESAs reviewed the topic of automation in financial advice in 2016.† One of the main benefits identified was the potentially lower costs of automated advice tools compared to traditional face-to-face advice, which could therefore contribute to make advice more affordable. Individuals could also benefit from access to a wider range of products and services. As far as the risks are concerned, issues identified included individuals potentially being exposed to unsuitable decisions because of lack of information about the process or limited opportunities to seek or provide clarifications and challenge decisions. Other risks include possible errors and/or functional limitations in the design of the algorithms that underpin the automated advice tools.

* Advice is understood here as defined by Article 2 (1) (15) of the IDD: the provision of a personal recommendation to a customer, either upon their request or at the initiative of the insurance distributor, in respect of one or more insurance contracts

† Report on automation in financial advice, Joint Committee of the ESAs, December 2016, https://esas-joint-committee.europa.eu/Publications/Reports/EBA%20BS%202016%20422%20(JC%20SC%20CPFI%20Final%20Report%20on%20automated%20advice%20tools).pdf

3.3. POST-SALE SERVICES AND ASSISTANCE

While the use of BDA in post-sales services and assistance is reportedly lower than in other stages of the insurance value chain (see Figure 11), many firms provided examples of how BDA is used in this stage. For example some firms have introduced automated answers in call centres and robotized customer service quality evaluation, including the use of speech analytics technologies to generate insights about issues such as complaint management, quality management, coaching of staff, fraud detection and consumer authentication.

Moreover, the above-mentioned concepts of CRM systems and Next Best Action are also relevant in this area of the value chain; firms explained how they use ML-based propensity predictive models or Natural Language Processing techniques to analyse both structured and unstructured data from customer's interactions to develop more meaningful, simpler and faster communications. They are also able to deliver more tailored services to the consumer, both online and offline, enhancing the consumer's satisfaction (commonly measured through indexes like the 'Net Promoter Score') and reduce churn rates.

VIRTUAL ASSISTANTS / CHATBOTS

An increasingly popular tool amongst insurance firms is the use of virtual assistants or chatbots, i.e. algorithms that enable 'human like' conversations with consumers via text or voice. Insurance firms use them predominantly in the post-sales and assistance area of the insurance value chain, to assist both customers as well as agents and brokers in many different ways.

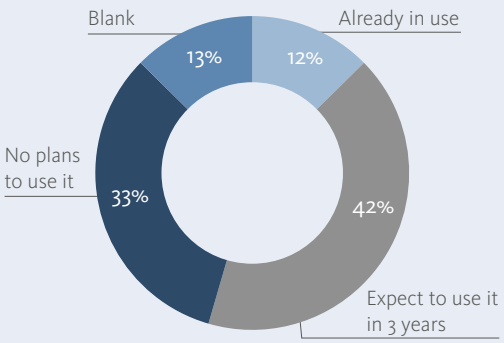
Chatbots are reportedly more sophisticated tools than traditional interactive voice response (IVR) tools already used by insurance firms in call centres for a long time. They are commonly embedded into messaging applications, available on a 24/7 basis, and they use natural language processing and other ML algorithms, which become increasingly efficient when trained with large quantities of data (e.g. behavioural data obtained from consumer interactions in the website or the mobile phone app).

Insurance firms use chatbots to support structuring and answering frequently asked questions ('FAQs') from consumers and to manage routine simple and non-sensitive servicing of the business. This includes the interaction with potential leads visiting the website, helping them to navigate the website and find information or collecting data to provide them a quote. They can also be used to guide consumers throughout the claims process (i.e. during the 'moment of truth').

The current and expected penetration of chatbots is bigger than the one of robo-advisors. As observed in Figure 15, 12% of the insurance firms use chatbots and 42% expect to use them within the next three years. One firm stated that they used a chabot in the past but they have stopped using it.

From those firms that already have a chatbot, 43% of them built the tool in-house, 39% bought the solution

Figure 15 – Chatbots penetration in insurance



Source: EIOPA BDA thematic review

off-the-shelf from a third-party provider, and 18% outsourced building the solution from a third-party provider. From those firms that are still not using chatbots but expect to use them within the next 3 years, a majority of them expect to deploy them with the involvement of third parties. Big Tech firms are amongst those third-party providers from whom insurance firms purchase chatbot solutions.

Chatbots present many benefits both for consumers and for the firms using them. On the one hand firms benefit from increased efficiency and reduced operational costs (e.g. one firm stated that within 3 years time they will expect to manage 50% of consumer queries through chatbots). On the other hand consumers benefit from user-friendly applications and 24/7 accesible services. Regarding the challenges, they are similar to the ones posed by robo-advisors, such as possible errors and/or functional limitations in the design of the algorithms or if consumers have limited opportunities to seek clarifications from humans.

Furthermore, insurance firms stated that BDA is also used to process the information collected via mobile phones apps which can include, depending on the app and the services provided to the consumer, a consumer's navigational information (e.g. appointments scheduled, reimbursements made, contacts made, number of clicks and behaviour on app elements), or device data (e.g. device type, operating system, usage hours), or geolocation and other types of telematics data. This information is then used by insurance firms in different ways, for instance to identify trends in app usage to improve the user's experience or to offer consumers a wide array of services linked to the connectivity and mobility of mobile phones devices. Many firms already count with more than one app.

Figure 16 – Mobile phone applications used by insurance firms

Types of services offered through mobile phone apps (depends on the app)	
First notice of loss (FNOL)	Submission of claims notifications
	Submission of photos of medical and pharmaceutical invoices or of car damages
	Monitoring of the claim's processing status
Sales	Sale of on-demand / short-term motor insurance policies
	Cross-selling of new insurance policies and/or renew and cancel existing ones
	Distribution of newsletters and promotions in selected shops
Risk mitigation and prevention	Premium discounts based driving behavior and/or healthy habits
	Coaching services including preventive push-notifications
Assistance	Automated emergency calls in case of accident
	Distance to the nearest branch, repair shop, gas station, parking, health provider etc. based on geolocation of the consumer
Consultations and appointments	Online video and chat medical consultations
	Make medical and dental appointments
	Consultation of consumer's medical records ("digital health card")
Policy information and contact details	Dashboard providing an overview of consumer's insurance policies
	Digitalized insurance contract's terms and conditions
	Contact details information of insurance branches, brokers, repair shops, medical partners etc.
	Submission of questions and complaints

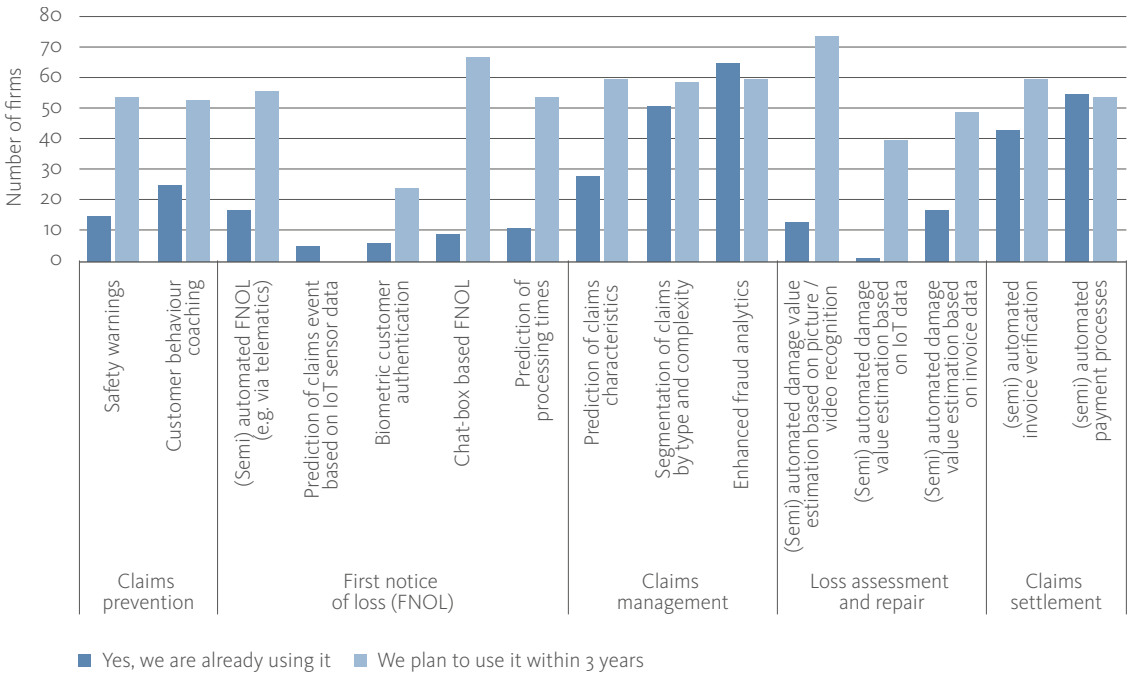
Source: EIOPA BDA thematic review

3.4. CLAIMS MANAGEMENT (INCLUDING FRAUD PREVENTION)

While some insurance firms still do not use BDA in claims management processes, the majority of firms already use BDA at some stage of the consumer claims journey and many more are planning to use them within the next three years, according to the survey. As it can be seen from Figure 17, BDA is currently used most often in fraud detection, followed by automated payment processes, segmentation of claims and invoice verification. However, none of the firms is currently using BDA in every stage of claims management, mostly due to the still low usage of IoT sensor data to predict claims.

Most insurance firms have leveraged on their in-house resources and their historical experience in processing claims to develop their own BDA tools used in this area of claims management (55% of total BDA tools/processes used in claims management). A minority of them have bought the tools off-the-shelf from third-party providers (25% of tools) and outsourced building the solution to a third-party provider (20% of tools). An overview of the penetration of the different BDA tools in claims management amongst European insurance firms is provided in Figure 17.

Figure 17 – Use of BDA in claims management



Source: EIOPA BDA thematic review, based on the classification of claims management processes from McKinsey&Company [14]

14 The classification of different processes is based on McKinsey&Company's claims consumer journey analysis; "Claims in the digital age: How insurance firms can get started", McKinsey&Company, April 2018, https://www.mckinsey.com/industries/financial-services/our-insights/claims-in-the-digital-age?reload

Claims prevention services like safety warning push notifications as well as consumer behaviour coaching refer to those services typically provided via mobile phone apps and other telematics devices which have already been commented in previous sections of this report.

Concerning First Notice of Loss (FNOL), firms explained how artificial neural networks could be used to predict the pathways for a claim that supports the claims process including the reserving process. For example, this can involve the prediction of the clinical milestones, processing times, setting quality baselines and workload distribution.

Firms using chatbots to guide consumers during the claims process emphasised that they are only involved to a degree where non-sensitive information is exchanged. Considering biometric customer authentication, which is a special category of personal data under the GDPR, some insurance firms described how they use fingerprints (e.g. using technology embedded in smart phones) and voice authentication analytics to identify consumers.

As far as the claims management stage is concerned, some firms use artificial neural networks to predict the claims characteristics and optimise their treatment. The latter includes for instance the identification of high-risk claims and claims with high recovery likelihoods in order to simplify analyses of hospital bills or to establish quantity controls. Claims can also be clustered depending on their complexity and fraud risk in order to automatically determine which team is responsible for the further processing of the claim.

USE OF BDA TO PREVENT FRAUD

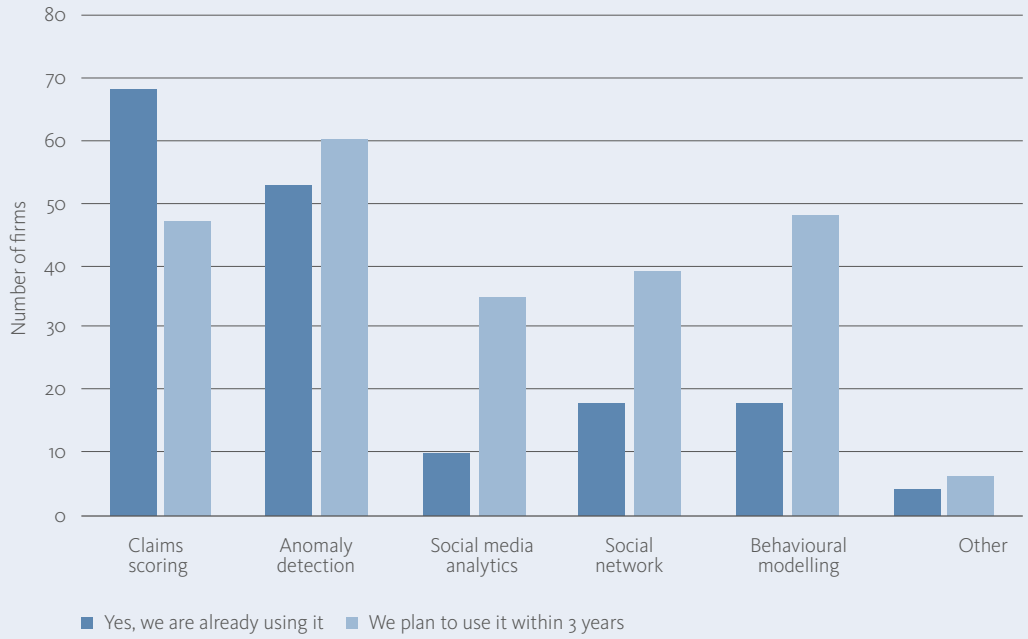
As shown in Figure 17, in claims management BDA is most often used to prevent fraud. Insurance fraud, i.e. intentionally bringing about an insurance event or causing the misconception of the occurrence of an insured event with the intention to receive insurance indemnity from the insurance firm, is a crime typified by the national law of the different Member States. According to Insurance Europe, the European insurance trade association, it is estimated to account for approximately 10% of all consumer claims.*

The expenses incurred by insurance firms in investigating and processing claims are known as loss adjustment expenses. Some insurance firms have special dedicated anti-fraud investigation units, often composed by personnel with a legal background as well as former police officers. In case of signs of consumer fraud, enhanced assessments are performed, which can include the use of private detectives. Insurance firms also commonly collaborate, creating claims and fraud databases within their respective national trade associations or in collaboration with public authorities.

Traditionally, there are two key stages in fraud-prevention: the first stage is prior to the conclusion of the contract; during the quotation process where insurance firms review the information provided by the consumer and cross-check it with internal and external sources of information such as fraud and claims databases or credit references. During the second phase, when processing claims, insurance firms' due diligence includes reviewing the documentation and evidence provided by the consumer to proof the loss and ensure the damages claimed by the consumer are accurate.**

BDA can support the detection of fraudulent claims in different ways. Most insurance firms have claims scoring tools, using ML algorithms in models trained to look for fraud patterns based on hundreds of different attributes (e.g. incident location, contract premium, number of previous claims by the policyholder etc.) and provide a fraud score for each claim. Often in combination with claims scoring techniques, insurance firms also use rule-based algorithms to assess claims, for instance by scanning invoices or images to automatically evaluate if the prices and damages are within the range of predefined/historical values or if they present anomalies. By flagging potentially fraudulent claims, investigators can focus on claims that are likely to be fraudulent and reduce the number of false positives and false negatives.

Figure 18 – Use of BDA in fraud detection



Source: EIOPA BDA thematic review, based on the classification of tools from Gartner†

Social media analytics, social network analytics and behavioural modelling are used less often amongst insurance firms. In this regard one firm stated that it assesses social media to analyse trends, although it does not really use BDA on it. Another firm described the generation of network diagrams in motor insurance, which are reviewed by fraud handlers alongside normal fraud referral processes, in order to help disclose hidden links between claims. Another firm stated that behavioural modelling is central to their health programme; it analyses different characteristics of health using BDA in order to best assess which behaviours best influence the overall health outcome.

* Insurance Europe, http://www.insuranceeurope.eu/fraud

** See EIOPA's fifth consumer trends report: https://europa.eu/!jc69pn

† Classification of types of BDA tools to prevent fraud is based on Gartner's analysis; Market Guide for Insurance Fraud Analytics, Gartner, 2016, https://www.gartner.com/doc/3241821/market-guide-insurance-fraud-analytics

Insurance firms also use BDA to analyse the information gathered via telematics devices (e.g. speed or G-forces) in order to estimate the severity of an accident and predict the characteristics of claims in motor insurance. Depending on the severity, the claims are either routed to undertaking's claims department or an emergency call is automatically alerted. According to some firms the segmentation of claims as well as the damage value estimation (reserve) can also be done by BDA tools processing the pictures and videos submitted by their customers.

Concerning claims settlements, several insurance firms use rule-based BDA tools to scan invoices from repair shops or medical partners to automatically evaluate if the prices are within the range of predefined values. This and other types of controls are automatically performed by BDA tools and then the payment transaction is automatically executed. However, in motor insurance some firms limit automated invoice and payment processes to glass repair claims and other small and simple claims. Also in some cases even though the process is automated, a clerk checks the invoice before payment is made.

Last but not least, a reduced number of insurance firms[15] stated that they use BDA for claims optimisation practices, i.e. the claims settlement offer provided to the consumer is influenced by BDA tools and processes estimating the likelihood or propensity that the consumer will accept or reject the claims settlement offer. This is reportedly done by introducing some demand analytics models to predict consumer behaviour into the calculation of the claims settlement offer. One firm stated that it outsources this service to a specialised third-party service provider. However, the majority of firms do not engage on these practices and state that their claims settlement offers are driven by objective facts (i.e. damage, cost for repair, etc.) and their customers expect them to be objective.

15 In total 35 firms stated that they use BDA for claims optimisation purposes, but based on their responses it is likely that many of these firms did not understand the question and therefore do not use BDA for these purposes. On the other hand, the responses of some of the firms indicated that they use such practices. There were also several blank responses to this question.

4. USE OF BDA FOR PRICING AND UNDERWRITING

In recent years, the European non-life insurance sector has experienced an increasingly competitive environment, in which insurance firms not only compete on services and cover offered, but also increasingly on price. This is the result of a wide range of factors such as the entrance of new competitors or consumers becoming more price sensitive (e.g. use of price comparison websites).[16]

In this context, insurance firms have started to adopt more sophisticated BDA-driven pricing models in order to optimize the profits with the help of the new possibilities offered by technological developments and new data sources. This has enabled a more granular segmentation of risks, increasing the effectiveness of risk selection, and allowing more risk-based pricing. This trend has also influenced the number and type of rating factors used by insurance firms in their pricing and underwriting models, both during the quoting process as well as on the renewal stage.

4.1. FINANCIAL INCLUSION/ EXCLUSION

According to the law of large numbers, the larger the number of units that are individually exposed to an event, the greater the likelihood that the actual results of that exposure will equal the expected results. This is used in insurance to explain the pooling of losses as an insurance mechanism; the larger the pool of resources of individuals, the higher predictability of the losses, which is reflected in the fact that the losses vary less around the average.

Consumers in insurance are grouped in pools with similar risk profiles (i.e. similar probabilities of making a claim), and premiums are based on the average risk across the pool. This has made it possible to set an adequate premium for each risk. There are cross-subsidies/risk balancing between those in the pool who do not make claims and those who suffer a loss. However, unlike in compulsory insurance or social security, insurance in its voluntary form does not seek cross-subsidies between members of the pool; the pooling of consumers and risk balancing between them are necessary in order to make the risks manageable/predictable.

Risk pools can vary in size; the ability to create bigger or smaller risk pools is based on the insurance firms' ability to distinguish the riskiness of different groups of consumers. It is argued that BDA improves the capacity of insurance firms to make a more granular segmentation of risks, increase the effectiveness of risk selection, and make pricing more risk-based. This micro-segmentation would result in smaller risk pools and a larger number of them, reflecting more accurately the risks of those within each pool.

This could potentially affect the ability of high-risk consumers to have access to affordable insurance coverage. EIOPA has collected a number of quantitative and qualitative indicators in an attempt to monitor the impact of BDA in the financial inclusion/exclusion of consumers in insurance. These indicators and their hypothetical evolution in a BDA context are presented in the table below.

16 Annex II of the present report shows that several firms offer consumers premiums a "street price" (i.e. premium paid by the consumer) lower than the "technical price" (i.e. the one calculated using traditional actuarial rating factors, such as expected claims costs, commissions, profit load and cost of capital).

Figure 19 – Financial inclusion/exclusion indicators

Indicator	Nature	Hypothesis in a BDA context
Standard deviation	Quantitative	The standard deviation of the average premium will increase
Insurability schemes	Quantitative	The number of members in the schemes will increase
Consumer complaints	Qualitative	The number of complaints of high-risk consumers will increase
Number of risk pools	Qualitative	The number of risk pools will increase
Rejection rates	Qualitative	The number of rejections of high-risk consumers will increase
Rating factors	Qualitative	The number of rating factors will increase

Source: EIOPA BDA thematic review

It is important to highlight that possible trends identified by these indicators may be explained by factors that are not related to BDA. For example, changes in the standard deviation may be affected by the competition dynamics or by inflation. It is therefore not possible to proof financial inclusion/exclusion issues by these indicators. Only if all or a number of the indicators pointed in the same direction they could potentially reveal signs (i.e. not proof) of the impact of BDA.

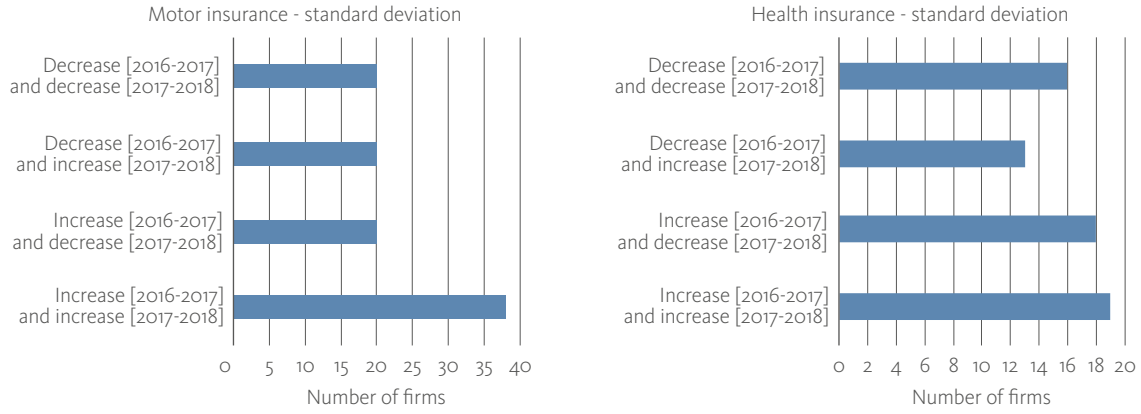
However, this is currently not the case based on the information available at EIOPA, although firms expect that the impact of BDA will be more pronounced in the years to come.

4.1.1. STANDARD DEVIATION

The standard deviation from the average premium measures the spread between the lower and the higher premiums (i.e. lower and higher risks pools, or lower and higher risk consumers). The hypothesis is that the standard deviation will increase over time because of BDA. EIOPA has collected from insurance firms the data on standard deviation in motor and health insurance lines of business for the years 2016, 2017, and 2018 (first half). The results are shown in the Figure 20 (motor insurance left and health insurance right).

There are more firms that do not show an increase in standard deviation in both consecutive years than firms that show an increase in both years. However, in motor insurance the group of firms that experience an increase in the standard deviation in both consecutive years is quite significant. Based on this information, and taking into account that other factors such as inflation or competition dynamics can also affect the evolution of the standard deviation, it is not possible to conclude that an increase in standard deviation is driven by BDA (see Annex 3 for further information).

Figure 20 – Evolution of the standard deviation of the average premium



Source: EIOPA BDA thematic review

4.1.2. INSURABILITY SCHEMES

Six countries[17] already have insurability schemes aiming to ensure that high-risk consumers have access to affordable motor insurance. While each scheme has its own specific rules, they are generally financed by the insurance industry in the concerned jurisdiction and provide motor third-party liability coverage to those consumers rejected by two or more insurance firms or who were offered a very high quote.

17 Spain (Consorcio de compensacion de seguros), the Netherlands (De Verennde), Luxembourg (Pool des risques agravées), Belgium (Bureau de tarification RC auto), Romania (Romanian Motor Insurance firms' Bureau (B.A.A.R.)) and Austria (Scheme for extraordinary risks)

Figure 21 – Membership evolution of national insurability schemes*

	2013	2014	2015	2016	2017	Growth 2013-2017
Austria	58	59	60	109	86	48%
Belgium	19,033	22,935	26,310	33,727	36,768	93%
Spain	66,969	58,159	55,266	54,025	54,169	-19%
Luxembourg	12	12	12	10	11	-8%
Romania	11	11	11	10	9	-18%
The Netherlands (GWP)	19,638,000	19,752,000	20,695,000	22,514,523	28,750,171	46%

Source: EIOPA BDA thematic review
* The data from Austria, Belgium, Spain, Luxembourg and Romania represents the evolution of the number of members in the respective insurability schemes. The data from the Netherlands represents the evolution of De Verennde's total GWP

Figure 21 shows that the number of members (represented by the evolution of GWP in the case of the Netherlands) has increased in three insurability schemes and decreased in three. Moreover, the number of members could partly be explained by the evolution of insurance contracts sales. It is also important to note that the number of members of insurability schemes represent a very small part of the population of the respective countries, and that members of the scheme commonly receive a renewal offer to stay within the scheme at the end of the contract.

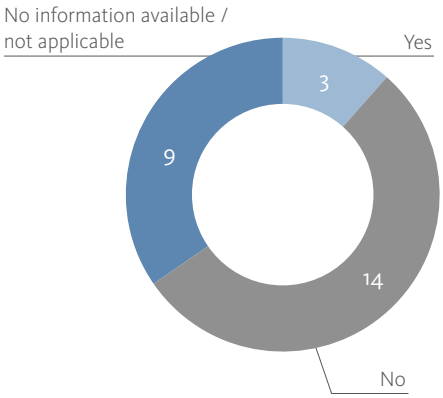
As far as health insurance is concerned, most EU Member States count with a public social security system that could potentially mitigate the effect of high-risk consumers being priced-out of the private health insurance market. In addition, in some private health insurance markets the law requires that rates are offered on a 'community' basis (as opposed to an 'individual' basis), which means that the question of whether or not to offer cover to a particular individual does not arise.

4.1.3. CONSUMER COMPLAINTS

The majority of national competent authorities (NCAs) from the insurance sector did not report any complaints related to BDA during 2017. The two consumer associations that provided input to EIOPA also did not identify complaints in this area. However, a significant number of NCAs expressed difficulties in providing an accurate statement given that the complaints that they receive are typically not categorised as BDA related or not. The argument is also made that consumers may not be aware that they are affected by BDA processes and therefore do not lodge a complaint for this purpose.

As far as those NCAs that reported BDA-related complaints, one NCA mentioned the case of consumers with disabilities having difficulties in accessing affordable health insurance. Another two NCAs mentioned the case of several consumers challenging the accuracy of a claims/anti-fraud database used by the insurance industry in those jurisdictions.

Figure 22 – BDA-related complaints

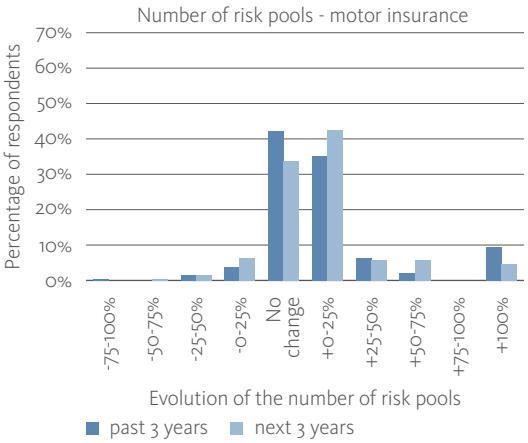


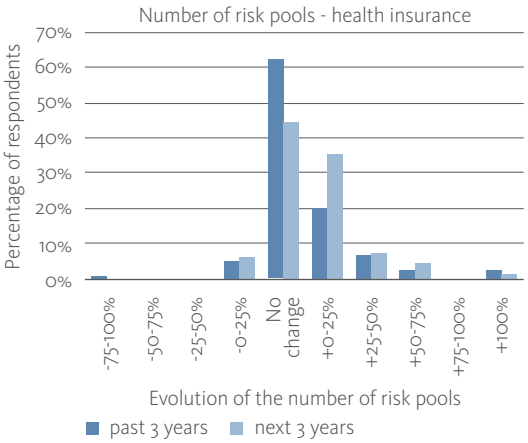
Source: EIOPA BDA thematic review

4.1.4. NUMBER OF RISK POOLS

Insurance firms were also asked about how the number of risk pools (i.e. homogenous risk groups) has evolved over the past three years, and they are expected to evolve over the next three years. For the past three years, a majority of respondents stated there has been no change in the number of risk pools in their organization in motor and health insurance. However, a significant proportion of firms responded there has been a slight increase in the number of risk pools at their organization in the last three years. This effect is more pronounced in motor insurance than in health insurance.

Figure 23 – Evolution of the number of risk pools*





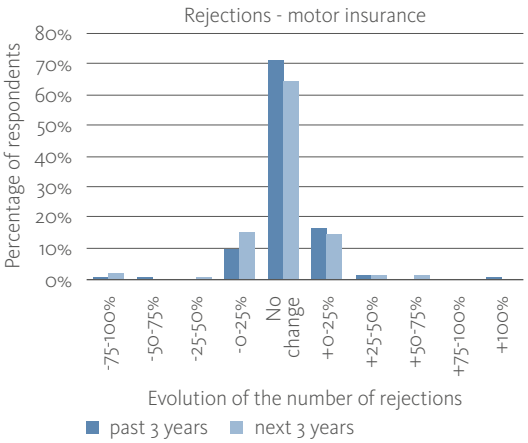
Source: EIOPA BDA thematic review
* Blank responses have not been considered in these graphics

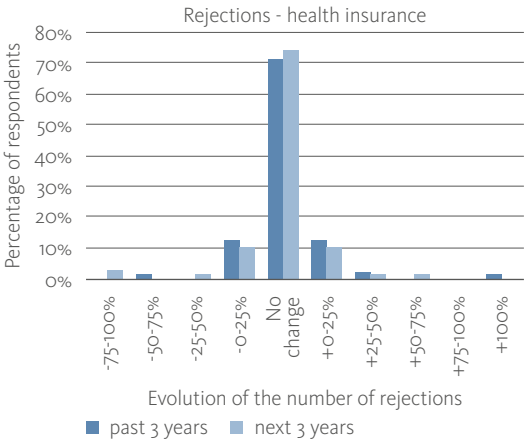
A majority of motor insurance firms expect an increase of between 0 – 25% in the number of risk pools in their organization over the next three years. A slightly smaller group believes there will be no change in the number of risk pools over the next 3 years. It is important to note that many firms in the motor insurance market commented that they have a very large number of individual tariffs and have done so for a long period of time (e.g. based on car model, year, etc.). In health insurance, a small majority of respondents is of the opinion that there will be no change in the number of risk pools over the next three years, although a slightly smaller group believe there will be an increase of between 0 – 25%.

4.1.5. REJECTION RATES

The significant majority of respondents stated that they have not experienced any changes in both past evolution and expected evolution of the number of rejections. Only a small number of firms expect the change in the number of rejections to slightly increase and slightly decrease over the next three years, across both motor and health. Among other reasons, the increase in the number or rejections could be partly explained by the enhancement of fraud detection enabled by BDA.

Figure 24 – Evolution of the number of rejections*





Source: EIOPA BDA thematic review
* Blank responses have not been considered in these graphics.

Interestingly, many firms commented on not rejecting consumers due to BDA, explaining that this is because they are required by law to provide a quote to all customers, with BDA improving their ability to calculate the 'right' price for each consumer. Others, mainly health insurance firms, mentioned that in their jurisdiction rates are offered on a 'community' basis, which means that the question of whether or not to offer cover to a particular individual does not arise.

Furthermore, some firms, particularly from the motor insurance sector, commented on the use of BDA enabling them to better differentiate between risks, and therefore being able to accept more customers and reducing the proportion of individuals for whom they refuse to offer cover. Telematics is mentioned frequently as a driver of better identification of high-risk drivers, with several firms making reference to the example of young drivers with no claims history being easier to price with richer data available.

In a number of cases insurance firms note that additional data could allow more accurate pricing but do not explicitly state that they will use this to expand into higher risk customer segments. Some also stated that insurance firms are highly risk averse and do not want to take on the greater volatility and risk of higher premium motor business even if they believe that BDA may enable them to price it more accurately.

Interestingly, one firm suggests that BDA will reduce prof-itability for high-risk drivers – presumably, because more insurance firms will be able to confidently offer a premi-um and so there will be greater competition in these seg-ments. Another firm states that innovation and medical advances are more material than BDA in allowing expan-sion of cover for health insurance.

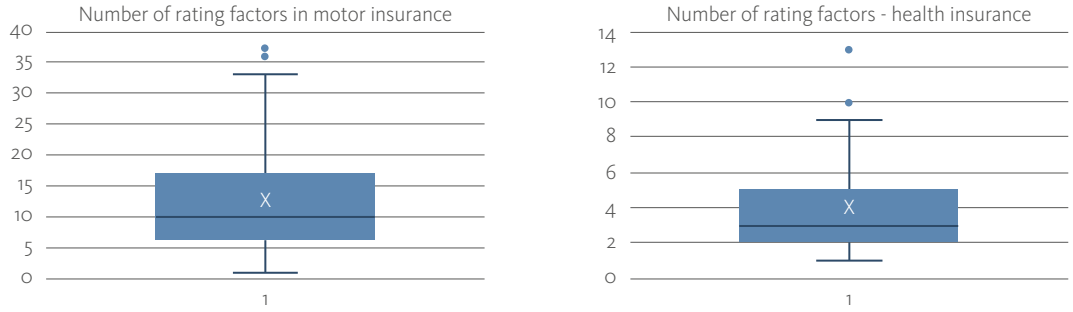
4.2. RATING FACTORS

The premium that consumers pay for their insurance policy depends on a number of individual characteristics assessed by insurance firms during the quote and renewal processes known as rating factors. Rating factors most often have a causal link and are used by insurance firms to measure the risk or probability that the individual will make a claim or suffer a loss, i.e. they are used to determine the risk profile of each consumer. Insurance firms consider the type, number and weight of the different rating factors used in their pricing models as part of their intellectual property.

Where direct evidence of risk (e.g. speed, braking, reaction to hazards for motor insurance) is not available, alternative rating factors can be used as a proxy. For example, in the absence of driving speed information, an insurance firm may use the type of car driven or number of speeding tickets by the policyholder as a proxy to infer this information; owners of sports cars and with a greater number of speeding tickets are expected to drive the fastest.

Based on the information from 128 motor insurance and 95 health insurance firms, the number of rating factors used in motor insurance is higher than those used in health insurance. Motor insurance firms use on average 13 rating factors, as opposed to an average of 4 in health insurance, with the median being 10 rating factors in motor insurance and 3 in health insurance.

Figure 25 - Number of rating factors used by insurance firms in 2018



Source: EIOPA BDA thematic review

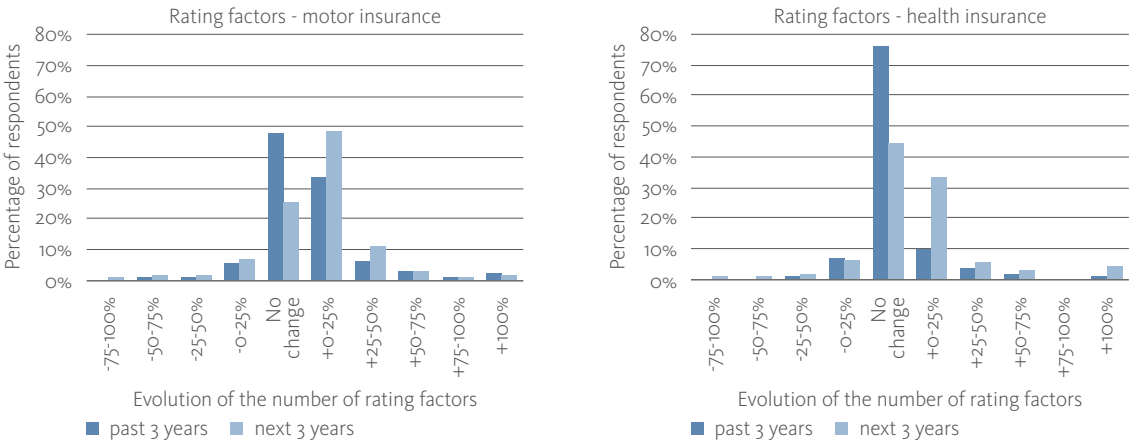
The spread or standard deviation of the number of rating factors in motor insurance is considerably higher than the one in health insurance; the number of rating factors used in motor insurance range from 1 or 2 amongst those firms prioritising smooth rating processes, to up to 37 rating factors used by those aiming to provide more accurate risk assessments. However, although not represented in Figure 25, EIOPA has received a detailed submission from a motor insurance firm using over 350 rating factors in their pricing model, using information from both internal and external sources. Notwithstanding the specificities of the business model of that firm, it shows the possibilities offered by the use of BDA in pricing and underwriting.

The largest proportion of respondents in both motor and health insurance markets state that there has been no change in the number of rating factors used in their pricing and underwriting processes in the past three years. However, around 1/3 of motor insurance firms state there has been an increase of between 0 – 25% in the number of rating factors used during this period. This compares to only 7% of respondents operating in the health insurance market.

The largest proportion of respondents in the motor insurance market believe that the number of rating factors used will increase by between 0 - 25% over the next three years. This differs in comparison to health insurance respondents, where the largest proportion believe that there will be no change, whilst a significant section of respondents do believe there will be an increase of between 0 - 25%. It is important to note that some respondents commented on refinements in use of existing rating factors. In addition, many intermediaries commented on the pricing being the sole responsibility of the underwriters. Finally, some respondents commented on the potential use of a range of factors from telematics data in the future.

As far as the type of rating factors used by insurance firms, EIOPA has aggregated the information received about motor insurance rating factors into 11 buckets that try to reflect the extent to which the underlying factor is related to risk, or could be perceived as correlated but with no evidence of causality, as indicated in Figure 27.

Figure 26 – Evolution of the number of rating factors*



Source: EIOPA BDA thematic review
* Blank responses have not been included in these graphics.

Figure 27 – Types of rating factors used in motor insurance

Motor insurance rating factors categories	Examples of rating factor included in this category	Influence on final premium (approximation)*	Type of information provided
Driver details	Age of driver, mileage, car usage	High	Perceived as having a direct causal link
Vehicle details	Horsepower, car model, car value	High	
Claims and experience	Bonus malus, year of obtaining the driving license	High	
Cover	Type of cover, deductibles	High	
Driver behaviour	Driving score, acceleration, telematics data	High	
Loyalty	Multi-subscription, renewal, tenure with company	Low	Perceived as having an indirect link to risk behaviours – more likely elasticity
Location	Postal code, region, area of residence	High	
Affluence	Credit score, kind of home ownership, occupation	Low	
Distribution	Sales or distribution channel	Low	
Non-risk (not captured)	E-mail address,** customer marketing opt-out preference, quote manipulation	Low	
Other	Miscellaneous	Low	Not clear, excluded

* Depends on concrete rating factor and firm. Therefore the aggregated value for the group is only an approximation taking into account the most repeated value of all the rating factors within that group (firms were asked to categorise each rating factor between High = is one of the 50% most influential rating factors they you use and low = is one of the 50% less influential factors that they use)

** Only one firm reported the use of this rating factor, reportedly to steer communications to paperless channels. However, it should also be noted that personal e-mail addresses might provide a lot of personal information about the consumer, since they often include the name, surname and age of the consumer, and further personal information can be derived therein. The media has recently reported examples of price discriminatory practices because of the use of email address as a rating factor.

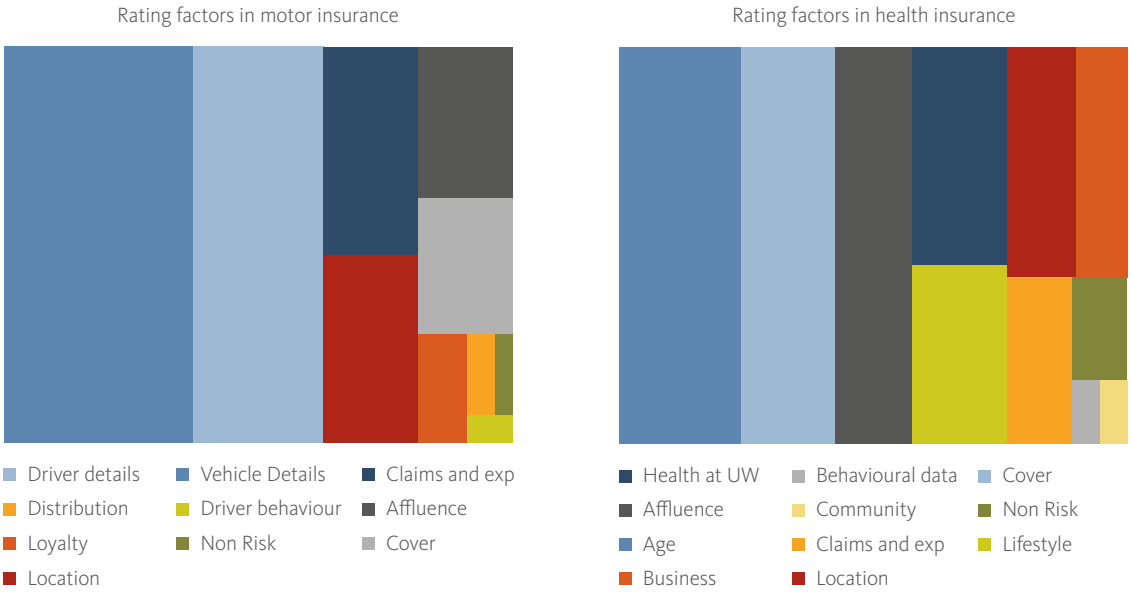
In health insurance the rating factors reported by firms have been grouped as indicated in Figure 28.

Figure 28 - Types of rating factors used in health insurance

Group	Examples of rating factor included in this category	Influence on final premium (approximation)	Type of information provided
Health at underwriting	Medical condition at the time of underwriting	High	Perceived as having a direct causal link
Age	Age of customer	High	
Behavioural data	Behavioural data	Low	
Claims and experience	Claims history	High	
Cover	Sum insured, number of people insured, deductibles	High	
Lifestyle	smoker, sports, dangerous activity, alcohol consumption	High	
Affluence	Profession, salary, payment periodicity, education	High	Perceived as having an indirect link to risk behaviours – more likely elasticity
Business	Number of employees, type of business	High	
Community	Community rating	High	
Location	Postal code, area of use, city of residence	High	
Non-risk	Competition, segmentation, sales channel	Low	
Other	Miscellaneous	Low	Not clear, excluded

Source: EIOPA BDA thematic review

Figure 29 - Rating factors used across the EU for motor and health insurance*



Source: EIOPA BDA thematic review

* Note: Category "Other" (circa 8% of the number of rating factors) has been excluded from the chart

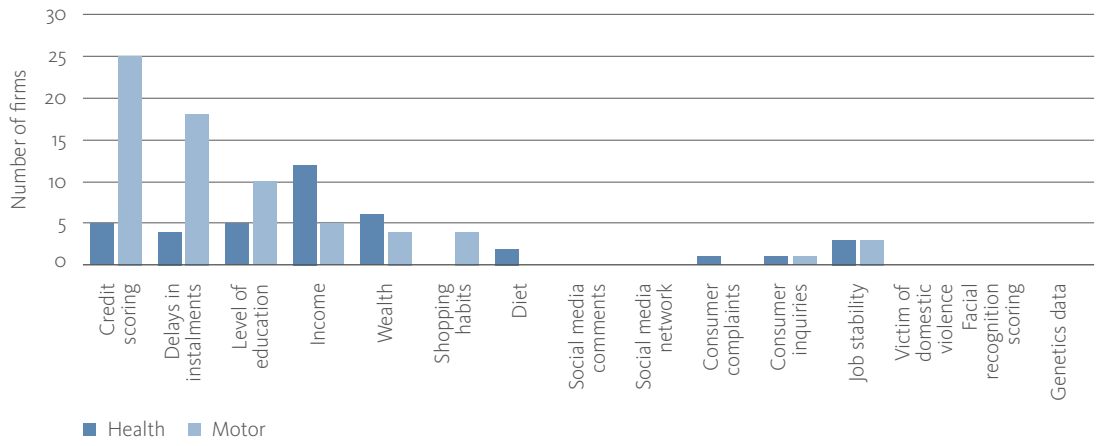
Based on these broad categories, the majority of rating factors used by motor insurance firms are related to driver details (e.g. age), followed by vehicle details (e.g. horsepower, type of vehicle etc.) and claims and experience (e.g. bonus malus). In health insurance, age is the rating factor more commonly used, followed by type of coverage (e.g. deductibles).

It is important to highlight that the above graphics show the distribution of the different groups of rating factors and not of individual rating factors nor their respective weight in the pricing model. Regarding the latter, the majority of motor insurance firms include the rating factors of age of the driver, address, claims experience and horsepower in their pricing models having a high influence in the final premium. In health insurance, age and address are the most common individual rating factors used and most often they have a high impact on the premium.

It should also be noted that the majority of rating factors used (80% in motor insurance and 67% in health insurance) are considered to have a direct causal link. This is the case of the categories such as driver and vehicle details in motor insurance or age and cover in health insurance. Other rating factors such as location or affluence (e.g. consumer's credit score or occupation), often have a high influence in the final premium and are perceived as having an indirect link to risk behaviours, more likely elasticity.

Moreover, to date very few firms use 'alternative' rating factors in determining the prices consumers pay, across both health and motor insurance. 'Income' is used most often by health insurance companies, but still relatively little across the industry. 'Credit scoring' and 'delays in instalments' are used with higher frequency in motor insurance, but still relatively little across the industry.

Figure 30 – Use of 'alternative' rating factors



Source: EIOPA BDA thematic review

Taking into account that 222 insurance firms participated in EIOPA's thematic review, the number of firms using data enrichment techniques to capture these rating factors is considerably low. However, some firms noted that although they may not use these specific rating factors, some of them can be inferred from other rating factors more commonly used by insurance firms.[18] For instance, a consumer's occupation or address are widely used rating factors that can be correlated with the level of education, income or wealth of the consumer, among other factors/characteristics.

4.3. MICRO-SEGMENTATION AND RENEWAL OFFERS

Whilst quite a few insurance firms declared that they will be able to use BDA to move towards individualised policy pricing, few of them have reached this level of sophistication yet. Some firms stated that they are already able to price at the level of the individual without using BDA. One firm explained this phenomenon as follows: 'As there are many more possible combinations of rating factors, there are a great number of different prices as well.'

Several firms gave some specific examples of the benefits that would arise from BDA in pricing and underwriting, citing the ability to more accurately identify fraud or that BDA would enhance the use of more accurate sociodemographic and financial variables as well as valuable behavioural information from web behaviour analytics. Several other firms mentioned the use of BDA for more accurate geographical data (i.e. substitution of traditional postal codes by more accurate 'micro-zoning'), and that it can support better cluster analysis.

A second tranche of firms believe that, whilst BDA will improve granularity of pricing and lead to more segmentation, it will not be possible (or desirable) to move to a segmentation of one. Some firms believe that fully individualised prices can only be achieved through direct monitoring of behaviour, namely via a telematics box for motor insurance or the use of wearable devices in health insurance. Moreover, a number of firms believe that the General Data Protection Regulation (GDPR) will effectively restrict the use of personal data and that it will be this restriction which holds back the development of individualised rating.

A third tranche of respondents link the issue of micro segmentation and individual pricing with price optimisation practices; they see BDA as a means of augmenting and enhancing, rather than replacing, existing pricing tech-

18 In addition, some of the firms that reported using these rating factors did not report them in the question where they were asked to provide all of the rating factors that they use. One possible explanation is that they understood the question about "alternative rating factors" as also including the inference of this information from other rating factors. Another possible explanation is that when they were asked to provide all of the rating factors used, they only provided the actuarial rating factors used to calculate the technical price (i.e. not the street price).

niques based on actuarial and statistical methods. They explain how, for some, micro segmentation is currently restricted to generalised linear models (GLM)[19] factors and curves, but tariff optimization using BDA like ML algorithms to build risk models and customer behaviour models (e.g. estimating propensity to churn or to shop around) can enable individual price optimization. This could be particularly relevant in motor insurance and at the renewal stage, where firms may attempt to compensate the initial pricing effort during the on-boarding phase in a highly competitive environment.

19 In brief, the generalized linear model (GLM) is a flexible generalization of ordinary linear regression that allows for response variables that have error distribution models other than a normal distribution. The GLM generalizes linear regression by allowing the linear model to be related to the response variable via a link function and by allowing the magnitude of the variance of each measurement to be a function of its predicted value.

PRICE OPTIMISATION PRACTICES

The expanding potential and use of BDA in the insurance sector may increase the ability of firms to identify opportunities to charge differential amounts to groups of consumers that are similar in terms of risk and cost to serve. As firms increase the range of data they consider and use more sophisticated analytical techniques in their pricing practices, they are better able to understand aspects such as consumers' price sensitivity and their likelihood to shop around and switch at point of renewal. This can increase the ability of insurance companies to use price optimisation practices when setting premiums.

Price optimisation refers to the practice of adjusting the premiums paid by different groups of consumers to achieve certain business objectives. Given an understanding of the behaviours and economic characteristics of consumers in the market, and an awareness of the behaviours of their competitors, insurance companies can aim to adjust premiums paid by different groups of consumers in ways unrelated to their risk or cost to serve to maximise overall profit. An increasing ability to present more tailored prices to consumers gives insurance companies the potential to become more competitive with consumers that are particularly price sensitive and consider switching often by lowering prices, and increase prices at renewal for other groups of consumers less likely to switch.

Consumers more prone to search for a better deal and switch at point of renewal are likely to benefit from price optimisation practices (or at least less likely to suffer any disadvantages) in comparison to consumers with similar risk and cost to serve characteristics, but a lower propensity to switch. On the other hand, consumers that are less price sensitive, less inclined to switch and more likely to renew their insurance products without searching for an alternative are more likely to lose out due to price optimisation. Insurance firms may identify that they are able to charge these consumers more than they would charge similar customers in terms of risk and cost to serve, and increase prices accordingly at point of renewal.

Price optimisation practices have drawn significant attention from regulators, industry, and commentators given the potential unfair treatment of some groups of consumers because of price optimisation. This could be particularly concerning where the groups of consumers that suffer most are more vulnerable consumers (e.g. old age, low income), or are suffering because of potentially unfair discriminatory practices. It may also be that consumers' vulnerability is what is causing them to suffer, for instance if they do not have time to search and switch to a cheaper provider due to a particular life circumstance they are in. Moreover, information such as postal Code (or the increasing use of more accurate 'micro-zoning') can discriminate against people in a poor area to obtain compulsory motor insurance, which may reinforce existing inequality.

In the USA the National Association of Insurance Commissioners (NAIC) published a White Paper analysing price optimisation and its use in insurance rate making, with a primary focus on personal lines in November 2015.* A number of states subsequently issued notices prohibiting or restricting the use of price optimisation or the concept of rating based on price elasticity in personal lines, indicating that price

optimisation results in rates that are unfairly discriminatory. In the UK the Financial Conduct Authority (FCA) recently conducted a thematic review on retail general insurance pricing approaches** and as a follow-up it has initiated a market study† to further explore pricing practices in general insurance markets (not only related to the use of BDA). The table below summarises six fundamental questions that the FCA considers when assessing fairness in the markets.

Figure 31 – The UK's Financial Conduct Authority approach to fairness

Who is harmed by price discrimination?	Wealthier consumers - eg time poor, cash rich	Consumers with characteristics which might be deemed vulnerable (eg low income, old ag, etc.)
How much are these individuals harmed?	Profitability difference between consumer segments is minimal and is immaterial to the harmed segment	Significant profitability differences and the harm has a significant adverse effect on the segment affected
How significant is the pool of people harmed?	Very small minority	Significant group of consumers
How are firms price discriminating?	Transparent and based on behaviour which consumers can easily change (eg switching)	Hidden and based on intrinsic characteristics which consumers cannot easily change (eg personal characteristics)
Is the product/service essential?	Product/service is considered non-essential but desired by some consumers	Essential product/service (eg current account or motor inssurance)
Does society view the price discrimination as egregious/socially unfair?	Little concern expressed about practices and firm behaviour widely accepted	Persistent and broad-based concern expressed and firm behaviour seen as poor conduct

Source: UK's Financial Conduct Authority††

The effect of price optimisation in European insurance markets and the extent to which it is taking place is unclear from the results of the EIOPA review and not necessary linked to the use of BDA but could be easier to conduct with BDA-tools. Based on the information provided to EIOPA, 59 firms already use or plan to use in the next 3 years BDA tools in pricing and underwriting. However only 19 of them made explicit reference to their use for price optimisation and/or churn models (the remaining 40 did not specify nor deny this purpose). Moreover, we do know from survey findings that the variety of data sources and range of BDA techniques to obtain greater insights into consumer behaviours in Europe, and with this comes the potential for an increase in the extent and sophistication of price optimisation practices. Indeed insurance firms will increasingly have a greater understanding of the demand elasticity and, more particularly, on the propensity of different consumers to churn, enabling them to adjust premiums accordingly.

* https://www.naic.org/documents/committees_c_catf_related_price_optimization_white_paper.pdf

** https://www.fca.org.uk/publication/thematic-reviews/tr18-4.pdf

† https://www.fca.org.uk/publication/market-studies/ms18-1-1.pdf

†† Price discrimination in financial services, Financial Conduct Authority, July 2018, https://www.fca.org.uk/publications/research/price-discrimination-financial-services

Many firms refer to differences between data used at the new business and renewal stages. However, it is often unclear whether or not the additional data used at renewal constitutes BDA. Moreover, a significant proportion (30%) of firms state that they do not use BDA at all or that there is no difference between the rating factors used to calculate new and renewal premium.

Six motor insurance firms specifically refer to price optimisation or elasticity analysis at the renewal stage, determined by using a retention or conversion model (e.g. amending the quoted rate on the basis of what the customer is likely to be willing to pay, rather than just on the probability that they will submit a claim). Although not explicitly stated by the firms responding to the survey, it is likely that these optimisation approaches make use of BDA.

A consistent theme at the renewal stage is that insurance firms acquire additional information about their customers that was not available at the new business stage. Common examples raised include claims experience, payment behaviour (e.g. timeliness of settling payments for those paying by instalment) and whether they have defaulted any payments or purchased any add-on products.

Firms also note that there are simple, inevitable differences between new and renewing customers (most obviously, age of both policyholder and (for motor insurance) vehicle). Furthermore some firms talk about offering loyalty discounts for renewing customers, and others about establishing caps and collars to the premium suggested by BDA tools in order to avoid a 'price shock' which would reduce the probability of successfully retaining the customer, and also potentially address unfair outcomes for vulnerable consumers.

5. BDA GOVERNANCE FRAMEWORK

There is a clear trend towards increasingly data-driven business models in insurance, with many insurance firms having developed their own BDA strategies or roadmaps or have included numerous BDA projects in their strategic business plans. This includes the development of 'enabler' tools aiming to improve the quality and quantity of datasets, governance policies, infrastructures, software and human resources skills necessary to deal with BDA.

Insurance firms that participated in EIOPA's thematic review were asked to explain specifically how they address potential data accuracy, fairness, and transparency issues arising from BDA. When responding to these questions, firms provided a detailed overview of the governance arrangements that they have regarding the use of BDA.

In this regard, firms acknowledged that mathematical calibration and validation of BDA models is a crucial and well-established step in the insurance sector. In particular Solvency II's governance requirements and the role played by the different key functions (i.e. audit, actuarial, compliance and risk management functions) were mentioned by many insurance undertakings as providing several 'lines of defence' to address potential issues arising from BDA.

Insurance firms also often referred to the new requirements introduced by the GDPR. In particular, firms often mentioned the requirement to appoint Data Protection Officers, develop data privacy impact assessments (DPIA) and respect key data protection principles such as the principle of accountability, data minimisation, data accuracy or data protection by default and by design.

In this regard, several firms stated that they are currently in the process of defining new data governance processes specifically for BDA. For example, some firms stated that they have recently appointed Chief Data Officers, others already have data governance policies approved by the Board of Directors and others have established data governance committees which regularly meet to discuss data-related issues. Firms also explained that external auditors periodically review their data governance processes.

Moreover, several insurance firms are currently developing new infrastructures/frameworks to implement BDA 'enablers' such as centralised data lakes/data warehouse/DataMart for data quality and data sanitisation. Using one primary data pool for various applications further ensures that different business users are doing data quality checks and providing feedback loops following the 'four eyes principle.' Furthermore, some firms have developed data inventories in order to evidence how they collect, share and use data in a lawful way.

Indeed, the firms that participated in the thematic review clarified that data governance processes are commonly done on a continuous basis and are not a one-off exercise; firms explained that new models such as the ones using ML algorithms commonly go through a training and testing phase and are reviewed by different departments before they are put into production. For example, to prevent discrimination or unfair outcomes of BDA processes, one firm described how it carefully reviews the impact of certain variables in the trained model before deciding whether to use them or not. Once the new model is introduced into the system, their performance is further monitored.

Finally, there might also be different data governance structures depending on the size of the organisation; one firm admitted that being a small size firm they do not have a department or group in the organisation to work specifically on this topic. Additionally, some firms explained that they only use anonymised or pseudo-anonymised information in their BDA processes or that they have considerably limited the use of non-anonymised data.

5.1. ACCURACY OF BDA INPUTS AND OUTPUTS

The majority of firms that participated in EIOPA's thematic review stated that they had not encountered any data accuracy issues or that they have robust governance processes to address them. Insurance firms reiterated that they do not use discriminatory variables in their BDA processes and that their BDA tools are designed in such a way that the output is unbiased and based only on statistical evidence. Some also explicitly mentioned that

they do not use variables that could suggest discriminatory characteristics through spurious correlations.

However, some firms admitted that it is challenging to find purchased data from third parties with the same quality standards than those datasets that they use internally. In this regard, some firms clarified that they only use internal data in their BDA processes. Others also only use external data from highly trustful sources such as public entities or data vendors that supply to the whole sector in a given jurisdiction.

Some firms only use data from third parties for marketing purposes where data accuracy is reportedly less relevant. Those firms using data from data vendors stated that they ensure its accuracy and lawfulness via specific clauses included in the contractual agreements signed with them in compliance with outsourcing legal requirements.

Concerning the accuracy of the BDA tools, many firms specifically stated that they do not use black box solutions that could potentially introduce discriminatory individual characteristics. Others only use these algorithms as support for their regular analytical work, not in core production processes, with all outputs being reviewed by (human) analysts before being used. Finally, some firms also explained that they do not directly manage BDA tools themselves; they outsource the use of these tools to third parties that often supply to several actors of the sector.

5.2. ETHICS AND FAIRNESS IN BDA

Insurance firms generally stated that they comply with existing legislation and therefore they had not identified any fairness/ethical issues. They explained that they do not collect sensitive data like gender, ethnic origin, religion etc. and some also specified that inputs or outputs that could lead to the inference/proxies of such characteristics were also not used. Furthermore, some firms stated that they have in place a number of internal governance tools such as internal codes of conduct, data privacy advisory panels or whistleblowing procedures to prevent discriminatory behaviours.

A reduced number of firms argued that their pricing and underwriting practices were fair because they do not use price optimisation practices where the premiums paid by the customer depends on factors other than their risk. In this regard, one firm explained that mutual insurance undertakings do not aim to maximise profits and therefore they do not use price optimisation practices in order to identify which consumers are willing to pay more.

Some firms acknowledged that ensuring fair and ethical outcomes could become increasingly challenging if black box ML algorithms like artificial neural networks were used in pricing and underwriting. Some insurance firms declared that they 'smoothed' the output of such algorithms, for instance by not using machine learning without human intervention or by establishing caps to the outputs of these tools in order to ensure ethical outcomes (e.g. not charging vulnerable customers excessively).

Regarding the potential difficulties to access insurance for high-risk consumers, several health insurance firms mentioned that public health systems or mandatory community rating existing in some jurisdictions should provide sufficient 'safety nets' in these cases. Motor insurance firms also referred to already existing mechanisms in some jurisdictions such as insurability schemes (see section 4.1.2) or the obligation of insurance firms to not reject motor third-party liability insurance (MTPL) consumers (albeit there is no limit in maximum premium). On the other hand, some firms considered that access issues cannot be solved by free competitive markets but rather need to be solved by public authorities.

There were mixed views regarding cross-subsidisation practices (e.g. charging new customers less than to existing customers); several respondents mentioned that they do not engage on these practices or that they are not allowed by their national legislation.[20] Others explained that this is justified because new customers usually have a higher risk than existing ones due to less information about their claims experience, as well as by competition dynamics and the effort made by insurance firms to attract new customers. Others linked the result of charging new customers less than to existing ones to price optimisation practices.

5.3. INFORMATION AND TRANSPARENCY

Insurance firms have to inform consumers about the types, sources and purposes of the use of personal data in

20 For example, Hungarian national insurance law forbids price differentiation between old and new customers in motor third-party liability insurance (MTPL) insurance.

their BDA processes in compliance with the requirements established on Article 13 and Article 14 of the GDPR.[21] In insurance, this is commonly done, even before the GDPR entered into force, via the terms and conditions of the insurance policy agreed by the consumer and/or through dedicated privacy notices. This information is also often made available on the websites of the firms. Firms admit that they can run to several pages in order to cover all of the requirements established by the GDPR.

In addition to explaining them how their data is processed and for what purposes, firms stated that they also inform consumers about the rights that they have with regards to processing of their personal data recognised in the GDPR. This includes rights such us the right to access, rectify, portability, restrict or erase their respective personal information. Consumers also are informed about the existence of automated decision-making processes, as well as of their right to object to such processes.

Regarding the latter, firms explained that they comply with the GDPR when they explain to their customers the outcomes of complex BDA processes. However, some firms acknowledged that if BDA tools such as machine-learning algorithms would be used for pricing and underwriting purposes, it would be very difficult to explain to consumers the outcome of such tools.

Finally, some insurance firms mentioned the legal requirement in their jurisdiction to publish the rating factors used in certain lines of business. Others mentioned that they voluntarily communicate to their customers the rating factors used, while others considered that such disclosure could adversely affect the rights and freedoms of insurance firms, including regarding trade secrets and intellectual property. On the other hand, the two consumer associations that provided input to EIOPA considered that the rating factors used by insurance firms should be published in order to allow consumers adopt informed decisions.

21 According to the GDPR, consumers have to be informed timely, appropriately and transparently, from their first touch-point with the company onwards, about all their personal data kept and about their rights in relation to their data. Upon legitimate request, consumers can at any time invoke their respective data subject rights e.g. to access, rectify, portability, restrict or erase (within legal boundaries) their respective personal data collected by the companies.

6. CONCLUSIONS AND NEXT STEPS

EIOPA is of the view that the thematic review has provided many concrete examples of potential benefits arising from BDA, both for consumers and for firms in the motor and health insurance sectors. EIOPA also acknowledges that insurance firms generally already have in place or are developing sound governance frameworks to mitigate the risks arising from BDA, although this thematic review did not assess the effectiveness of such governance frameworks.

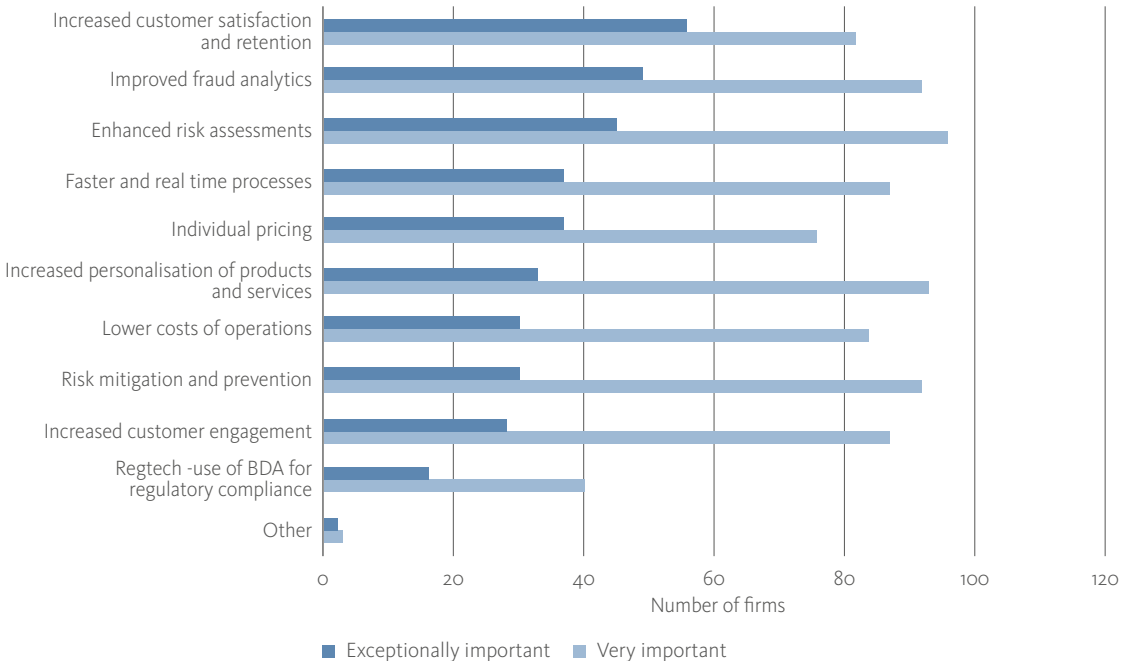
While there is already a comprehensive regulatory framework in this area (see Annex 3 for further details), EIOPA considers that there are risks arising from BDA that need to be further addressed in practice. Some of these risks are not new but they are amplified in the context of BDA. This is particularly the case of ethics and fairness issues in the use of BDA, as well as regarding the accuracy, transparency, auditability, and explainability of BDA tools such as AI and ML algorithms.

6.1. OPPORTUNITIES

BDA enables firms to better understand the needs, characteristics and lifestyles of consumers enabling them to develop more accurate risks assessments. This also allows firms to develop more personalised and convenient products and services for consumers; the fact that they can be delivered in an increasingly tailored and timely manner reportedly improves consumer's engagement and user experience.

In the context of the Internet of Things (IoT), the insurance sector has seen the emergence of usage-based insurance products. While it is still at an early stage of development, motor and health insurance customers already can obtain a more accurate calculation of their insurance premium based on the driving information collected through telematics devices installed in their cars or through health wearables. Some firms also use these telematics devices

Figure 32 – BDA opportunities according to insurance firms



Source: EIOPA BDA thematic review

to offer consumers a number of risk prevention and mitigation services such as coaching services or automatic assistance services in case of accident.

The use of BDA allows firms to develop more accurate risk assessments and new rating factors, which can be used to introduce new products for specific targets, markets and groups of coverage where previously was not possible. Among other things, this can lead to the financial inclusion of certain groups of consumers which were previously excluded. For example, young drivers with limited driving experience reportedly have access to more affordable motor insurance if they install telematics devices in their car. In addition, the accuracy and objectivity of the calculation of technical provisions can also be enhanced by using BDA.

From a sales and distribution perspective, the development of CRM systems incorporating all the information from consumers into one single platform allows firms to develop increasingly personalised and targeted marketing campaigns. Also interesting is the development of the 'Next Best Action' approach, where BDA tools such as ML can be used by firms to develop more consumer-centric cross-selling and up-selling models, for instance following a "consumers that bought this might also buy this" approach.

The penetration of robo-advisors could potentially allow consumers to access more affordable advice. However, the level of penetration of these tools is still limited compared to chatbots using natural language processing and other ML algorithms. Chatbots are increasingly popular in view of the wide range of possibilities that they offer for servicing the customer in simple and non-sensitive procedures (e.g. answering Q&As or guiding consumers through the quoting process). Their availability on a 24/7 basis and the possibility to use them from any location are reportedly seen as convenient services by consumers.

EIOPA believes that one key development in the area of BDA is the increasing use of mobile phone technology to collect new sources of data and interact with consumers; in particular, the thematic review has gathered detailed examples of how insurance firms provide different types of services to their customers through mobile phone applications. For example, consumers can submit claims (attaching pictures of invoices or car damage), or buy short-term/on-demand motor insurance policies, or make medical and dental appointments via their mobile phone apps.

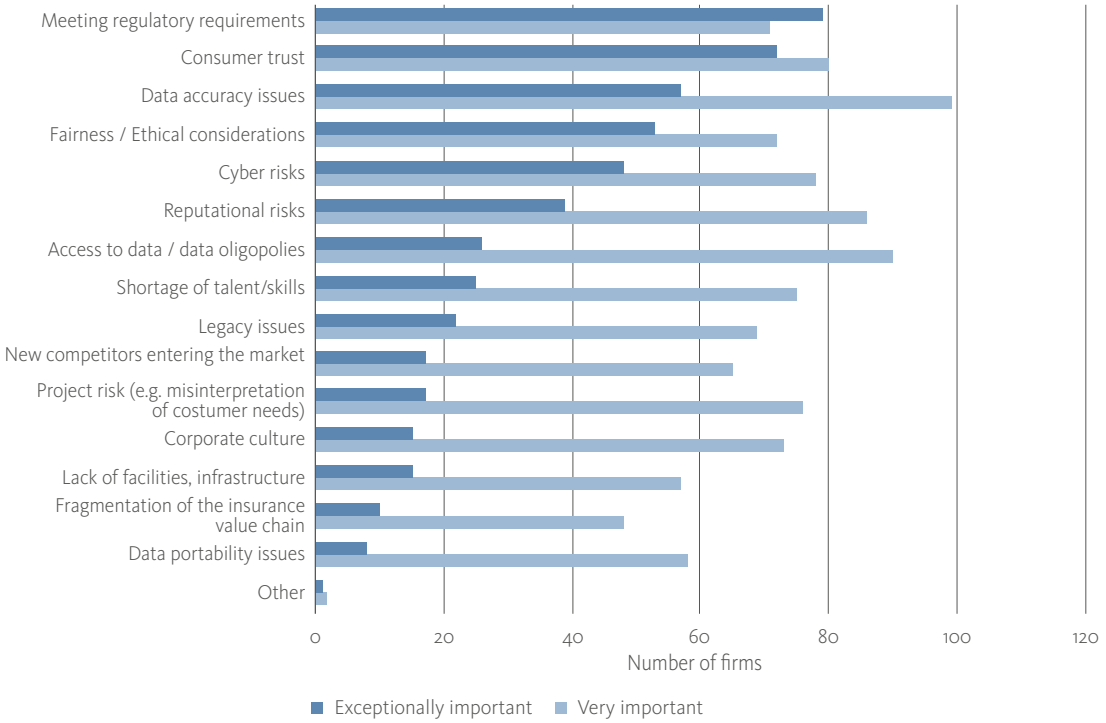
In the area of claims management, BDA can support the detection of fraudulent claims in different ways. Most insurance firms use claims scoring and anomaly detection tools, where tools such as ML algorithms are trained to look for fraud patterns based on hundreds of different attributes (e.g. incident location, contract premium, number of previous claims by the policyholder etc.). By flagging potentially fraudulent claims, investigators can focus on claims that are likely to be fraudulent and reduce the number of false positives.

6.2. CHALLENGES

As noted in the Declaration on Ethics and Data Protection in Artificial Intelligence during the 40th International Conference of Data Protection and Privacy Commissioners that took place in Brussels on October 2018,[22] some datasets used in BDA processes can be biased and mask some forms of prohibited discriminations. Given that BDA tools such as AI and ML algorithms rely on historic data for 'training', any biases (e.g. societal / ethical) inherent in the historic data will be reflected in the output of these algorithms.

22 Declaration on Ethics and Data Protection in Artificial Intelligence, 40th International Conference of Data Protection and Privacy Commissioners, October 2018, https://icdppc.org/wp-content/uploads/2018/10/20180922_ICDPPC-40th_AI-Declaration_ADOPTED.pdf

Figure 33 – BDA challenges according to insurance firms



Source: EIOPA BDA thematic review

This issue becomes more significant where specific judgements of a (black box) algorithm cannot be specifically explained in a meaningful way, raising fundamental questions about the accountability of those firms using them. Arguably, this is less relevant in areas such as marketing campaigns, where BDA are mainly used to launch more tailored communication campaigns and personalised offers to consumers. However, in the context of pricing and underwriting, if algorithms rely on biased datasets or rating factors,[23] this could potentially result in illegal price discrimination practices if not handled with the adequate due diligence and in accordance with generally accepted actuarial principles. Furthermore, if the output is based on correlations which are falsely assumed to be causations, then the decision-making process would be biased as well.

Additionally, an algorithm that is not sufficiently explainable, transparent, auditable or accurate can jeopardize the overall solvency position of an insurance undertaking. For example if the technical provisions or prices are calculated incorrectly, and due to a black box approach as well as a lack of internal control mechanisms this situation stays unnoticed, this can negatively impact the insurance undertaking's solvency.

EIOPA also considers that current practices such as price optimisation practices, i.e. when the premiums paid by consumers not only depends on their risk but also on their price sensitivity and their likelihood to shop around, could also be challenging from a fairness and ethical perspective. This would be particularly important regarding their potentially adverse impact on vulnerable consumers (e.g. old age, low income, low level of studies). This could also be the case of claims optimisation practices, where the compensation paid to the consumer suffering a loss does not only depend on objective facts like the damage, cost for repair, medical expenses etc. The potential mismatch between consumer's expectations and the actual practices could have a negative impact on consumer's trust, which could eventually become a challenge to the stability of the insurance sector.

23 From the almost 1000 rating factors analysed, only a handful of rating factors were considered as having a high risks of introducing illegal price discrimination criteria (e.g. rating factors based on the email address or the nationality of the consumer). EIOPA will follow-up with the relevant NCAs and ask them to address these issues with the firms concerned.

Moreover, the use of new data sources such as genetics data, credit card and bank account data, or other types of behavioural and sociodemographic data, both from internal and external sources, also need to be carefully monitored. In particular, datasets from external sources and those reflecting lifestyles of consumers (e.g. shopping habits, bank account and credit card data) are perceived as having greater risks of facing accuracy issues, or providing correlations with prohibited criteria, and therefore need to be handled with special care.[24]

In addition, the use of genetics data for pricing and underwriting in health insurance, which is currently in practice non-existent in the European insurance sector based on the responses of the thematic review, could also potentially lead to exclusion issues of high-risk consumers. Concerning the latter, EIOPA's thematic review did not find evidence that increased granularity of risks assessments is causing exclusion issues for high-risk consumers, both in health and in motor insurance. Furthermore, existing institutions in certain Member States such as insurability schemes, public health systems, or community rating, may well, in those specific Member States where they exist, provide sufficient 'safety nets' for the time being.

EIOPA notes that the issue of information and transparency vis-à-vis consumers concerning the types and sources of personal data used by insurance firms is already extensively covered by GDPR transparency requirements. It is nevertheless questionable and may deserve a closer consideration whether consumers are fully aware of how their personal data is being used when they accept the terms and conditions of their insurance policy and/or dedicated privacy notices, in particular if these documents run to several pages long and are not comparable between different firms. It is also debatable how firms can meet GDPR's requirement to explain to consumers in a meaningful way the functioning of BDA tools in the context of 'black box' algorithms.

The increasing exposure to cyber security risks is also seen as a major risk by EIOPA, noting that this can also be considered as an opportunity for the insurance sector via the commercialisation of cyber insurance policies.[25] The increasing use of cloud computing services could potentially also raise some data security issues and concentration risks for insurance firms. EIOPA acknowledges that the increasing outsourcing activity in the sector allows insurance firms to improve the efficiency of their internal processes and obtain quick access to new technologies and business models. However, an excessive concentration in the number of providers in certain strategic services/technologies can potentially disrupt the efficient functioning of value chains, leading to situations of 'reverse outsourcing,' i.e. the reversal of the traditional power relationship between insurance firms and the subcontractor. The latter could potentially undermine the effectiveness of the rules that today govern the outsourcing of 'essential services'.

6.3. NEXT STEPS

The speed of innovation developments, competition dynamics in the markets and firm's strategic business plans indicate that although the use of new data sources and the adoption of BDA tools in the insurance sector may not be widespread yet, they are certainly expected to significantly increase in the next three years. This has already drawn considerable attention from regulators,[26] industry,[27] and commentators, given the numerous potential benefits for society arising from BDA, but also given the

24 This is the reasoning behind the circular issued by the NY insurance State Authority in January 2019 cautioning insurance firms of the use of external data sources in life insurance due to possible biases in the information https://www.dfs.ny.gov/industry_guidance/circular_letters/cl2019_01

25 Understanding Cyber Insurance - A Structured Dialogue with Insurance Companies, EIOPA, August 2018, https://eiopa.europa.eu/Pages/News/Deeper-understanding-of-cyber-risk-needed-%E2%80%93-A-core-challenge-for-the-European-Insurance-Industry.aspx

26 In Europe, UK's FCA Feedback Statement on Call for Inputs on Big Data in retail general insurance, November 2016, https://www.fca.org.uk/publications/feedback-statements/fs16-5-call-inputs-big-data-retail-general-insurance, Germany's supervisor Bafin published a study "Big Data meets artificial intelligence" in July 2018 (https://www.bafin.de/SharedDocs/Downloads/EN/dl_bdai_studie_en.html). France's ACPR also published a discussion paper "Artificial intelligence: challenges for the financial sector" in December 2018 (https://acpr.banque-france.fr/en/artificial-intelligence-challenges-financial-sector). The European Commission has also established a High Level Expert Group on Artificial Intelligence, which published draft ethics guidelines for trustworthy AI on December 2018 (https://ec.europa.eu/digital-single-market/en/news/have-your-say-european-expert-group-seeks-feedback-draft-ethics-guidelines-trustworthy). The EU Agency for Fundamental Rights has also published a report titled, "BigData: Discrimination in data-supported decision making", May 2018, https://fra.europa.eu/en/publication/2018/big-data-discrimination. The Council of Europe in relation to possible standard setting instruments concerning AI: https://www.coe.int/en/web/freedom-expression/msi-aut

27 The French Actuary Association has developed some guidance in the area of Big Data in its "Norme de pratique relative à l'utilisation et la protection des données massives des données personnelles", Novembre 2017, (https://www.institutdesactuaires.com/decouvrir-l-institut/textes/normes-professionnelles-25) and the Association of British Insurance firms recently reached a compromise to address the issue of excessive differences between new customer premiums and subsequent renewal premiums that unfairly penalise long-standing customers, May 2018, (https://www.abi.org.uk/news/news-articles/2018/05/insurance-industry-takes-action-on-excessive-differences-between-new-customer-premiums-and-renewals/)

potential for some groups of consumers to be worse off with these practices without sound governance frameworks in place.

As a follow of the thematic review, EIOPA will further assess the issue of supervision of AI/ML black box algorithms. On the one hand the Joint Committee of the ESAs will look into the topic of artificial intelligence from a cross-sectorial perspective in 2019. In parallel, EIOPA will seek to further assess how AI and ML can be best supervised in practice, including assessing how their supervision differs from other models commonly used in insurance such as generalised linear modelling (GLM), generalised Bayesian model or decision trees.

In this regard, different options will be considered, such as introducing specific governance requirements for specific BDA tools and algorithms. This could include reviewing the role that Solvency II's key functions (and in particular the actuarial function) should play in this context. Other options include promoting enhanced transparency, explainability and auditability of algorithms, and/or restrictions on inputs and/or outputs and/or human intervention requirements, particularly in those areas where BDA tools might have a significant impact on consumers. Different capacity-building activities such as workshops or seminars for NCAs will also be promoted.

EIOPA will also discuss with the industry, consumer associations, academia and other interested stakeholders the issue of ethics and fairness regarding the use BDA in insurance. EIOPA acknowledges that different stakeholders have called for greater clarity about what are the supervisory expectations in this area, so EIOPA will initiate a debate on ethics and fairness in BDA, to explore whether further convergence and consistency from a sectorally specific perspective is needed.

Moreover, in the context of the EU-US insurance dialogue,[28] EIOPA will further explore third-party vendor issues; their current regulatory oversight and whether or how this framework focuses on issues surrounding BDA accuracy concerns, and the extent to which the current regulatory perimeter is addressing the ability for regulatory oversight. Possible disclosures to applicants and policyholders specifically about how rating factors and third- party reports are used by insurance firms will also be discussed.

In the area of outsourcing, EIOPA will issue guidelines on outsourcing of cloud computing services by insurance undertakings in order to promote supervisory convergence amongst NCAs and transparency considering regulatory requirements. EIOPA will also initiate a new workstream on new business models and ecosystems arising from InsurTech.

Finally, EIOPA will continue its ongoing work on cyber insurance and cyber security. This includes joint advice from the ESAs to the European Commission on cyber resilience testing. The ESAs will also develop joint advice to the European Commission on possible legislative improvements in the area of cyber and IT security. In addition, EIOPA is currently developing a report using granular data from the 2018 insurance stress test to assess the vulnerabilities of insurance firms regarding cyber risks and also the challenges and risks involving cyber underwriting.

28 The EU-U.S. Insurance Dialogue Project (EU-U.S. Project) began in early 2012, as an initiative by the European Commission, the European Insurance and Occupational Pensions Authority (EIOPA), the Federal Insurance Office of the U.S. Department of Treasury (FIO), and the National Association of Insurance Commissioners (NAIC) to enhance mutual understanding and cooperation between the European Union (EU) and the United States for the benefit of insurance consumers, business opportunity, and effective supervision. In 2018, the EU-U.S. Project's members continued the work focusing on the use of big data besides the other focus areas relating to cybersecurity risk, the cyber insurance market and intra-group transactions. A public forum was held on November 2018, https://eiopa.europa.eu/Pages/Events/PUBLIC-FORUM-EU-US-INSURANCE-PROJECT.aspx

ANNEX I: METHODOLOGY

On 15 March 2018 the Joint Committee of the ESAs published its key findings of its cross-sectorial review of the use of Big Data by financial institutions.[29] The report highlighted that there are a wide array of potential benefits arising from Big Data. However, a number of risks were also identified requiring more in-depth analysis and supervisory oversight.

In this context, EIOPA's Board of Supervisors decided to launch a thematic review on the use of Big Data Analytics specifically by insurance firms. The aim of the thematic review was to gather further empirical evidence on the benefits and risks arising from BDA in motor and health insurance lines of business.[30] It was decided to focus on these two lines of business in order to keep the exercise proportionate and because the potential impact of BDA in these two lines of business was reportedly high. The thematic review was officially launched during the summer of 2018.

The questionnaires that EIOPA circulated to national competent authorities, consumer associations and insurance firms can be consulted in EIOPA's Website.[31] The responses are not published due to the commercially sensitive nature of the information gathered. The questionnaires essentially consisted of a number of qualitative and quantitative questions, which gathered all the necessary evidence from different types of sources that would allow EIOPA to ascertain to which extent BDA risks and benefits were substantiated or not and/or whether they would require further regulatory or supervisory action.

For this purpose, each NCA was asked to distribute the industry questionnaire to insurance undertakings that covered at least 60% of the market share in each of their respective motor and health insurance markets. They were also asked to provide data from at least two start-ups[32] (if they existed in their jurisdictions). Provided that the they had a mandate to supervise insurance intermediaries, NCAs were also asked to circulate the survey amongst four insurance intermediaries, ideally combining small and bigger size entities, as well as innovative ones with those that count with more traditional business models.

Once they had collected the data from their respective jurisdictions, NCAs were asked to submit the responses anonymised to EIOPA. The NCA and consumer association surveys were distributed and collected directly by EIOPA. In the case of the consumer association surveys, EIOPA contacted BEUC and Better Finance and asked them to distribute the surveys through their respective members (i.e. national consumer associations).

In total 222 firms from the insurance sector from 28 Member States replied to EIOPA's questionnaire.[33] In addition, 24 NCAs and two national consumer associations (from the Netherlands and Portugal) submitted a response to EIOPA. This report is essentially descriptive of the information gathered from these stakeholders. In addition, external sources of information such as research studies from private and public entities have been used (and referenced) to provide some background information about the topics described in the blue-boxes. The conclusions section of the report represents EIOPA's views.

Further details about the firms that participated in the thematic review are provided below:

29 Report on the use of Big Data by financial institutions, Joint Committee of the ESAs, 15 March 2018, https://eiopa.europa.eu/Publications/Other%20Documents/JC-2018-04%20Joint%20Committee%20Final%20Report%20on%20Big%20Data.pdf

30 Due to specificities of the Irish health insurance market, Irish firms provided input based on the "medical expense insurance" line of business instead of the "health insurance" lines of business.

31 The questionnaires that EIOPA circulated to national competent authorities, consumer associations and insurance firms can be consulted in EIOPA's Website: https://eiopa.europa.eu/publications/other-documents

32 For the purpose of the distribution of the industry questionnaire, start-ups were defined as firms that have been granted an insurance license within the last 10 years, which commonly are SMEs and count with innovative digital business models. The latter may also include small autonomous subsidiaries that form part of larger insurance groups.

33 Participating countries include the following: Austria, Belgium, Bulgaria, Croatia, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, the Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, the United Kingdom, Liechtenstein and Norway. Malta, Cyprus and Island did not provide any input from the insurance industry in their markets

Figure 34 – Overview of participating insurance firms

Total participants	
Member States	28
Responses	222

Incumbent versus start-up	
Incumbents	170
Start-ups (year 2010+)	20
Unknown year	32

Type of license/authorisation	
Insurance undertaking license	134
Insurance intermediary license	50
Both	25
Unknown license/authorisation	13

Line of business	
Motor insurance	69
Health insurance	35
Both	105
Unknown line of business	13

Source: EIOPA BDA thematic review

A multidisciplinary workstream with representatives from NCAs and EIOPA was created in order to analyse all the information gathered through the BDA thematic review. During the analysis of the information, the workstream encountered a number of data quality issues such as incomplete or inconsistent responses. However, most of the submissions were on average of good quality and provided very valuable information about the impact of BDA in motor and health insurance.

Not surprisingly, the quantitative part of the industry survey was the one where the responses were of lesser quality; some firms had difficulties obtaining the information requested and others argued that the question was not sufficiently granular to obtain useful information from it. Therefore, the use of this quantitative information in the conclusions of the thematic review have been limited (for further details see Annex 2).

Moreover, some firms were reluctant to provide the information about rating factors used in their pricing model, arguing that this information was part of the firm's intellectual property. EIOPA had tried to mitigate this by not requesting the exact weight of each of the rating factors used in the models, but rather asking the firms to classify them into two broad categories (High = is one of the 50% most influential rating factors they use; Low = is one of the 50% less influential factors that they use).

The results of the analysis of the information gathered in the thematic review were presented and discussed before EIOPA's InsurTech Task Force. EIOPA has also discussed the findings with the European Data Protection Supervisor. The report was ultimately approved by EIOPA's Board of Supervisors.

ANNEX II: ANALYSING THE IMPACT OF BDA IN PRICE DIFFERENTIATION

EIOPA's BDA industry questionnaire included one quantitative questions designed to assess whether pricing between policyholders is becoming more differentiated over time (both from an assessment of the underlying risk as well as the final price paid by the consumer). The idea was to find evidence of the use of BDA by using information regarding average premium and standard deviation in combination with other qualitative indicators such as number of risk-pools or number of rating-factors.

Some of these indicators were similar to the ones used by the 'solidarity monitor' developed by Dutch Insurance Association in 2017, which monitors the impact of BDA in insurance pricing in the Netherlands.[34] However, the information gathered by EIOPA had a different perspective than the Dutch solidary monitor and insurance firms were asked to provide data on a less granular level (in particular EIOPA did not collect data on 'representative persons' given that they would likely not be comparable across the different EU Member States).

It is important to highlight that, similar to the solidarity monitor, the data collected by EIOPA does not reveal whether the potential price differentiation is caused by BDA or by something else. For example, changes in the standard deviation may be affected by the competition dynamics or by inflation. It is therefore not possible to proof financial inclusion /exclusion issues by these indicators. Only if all the indicators pointed in the same direction they could potentially reveal signs (i.e. not proof) of the impact of BDA.

34 Verbond lanceert eerste Solidariteitsmonitor, Verbond van Verzekeraars, October 2017

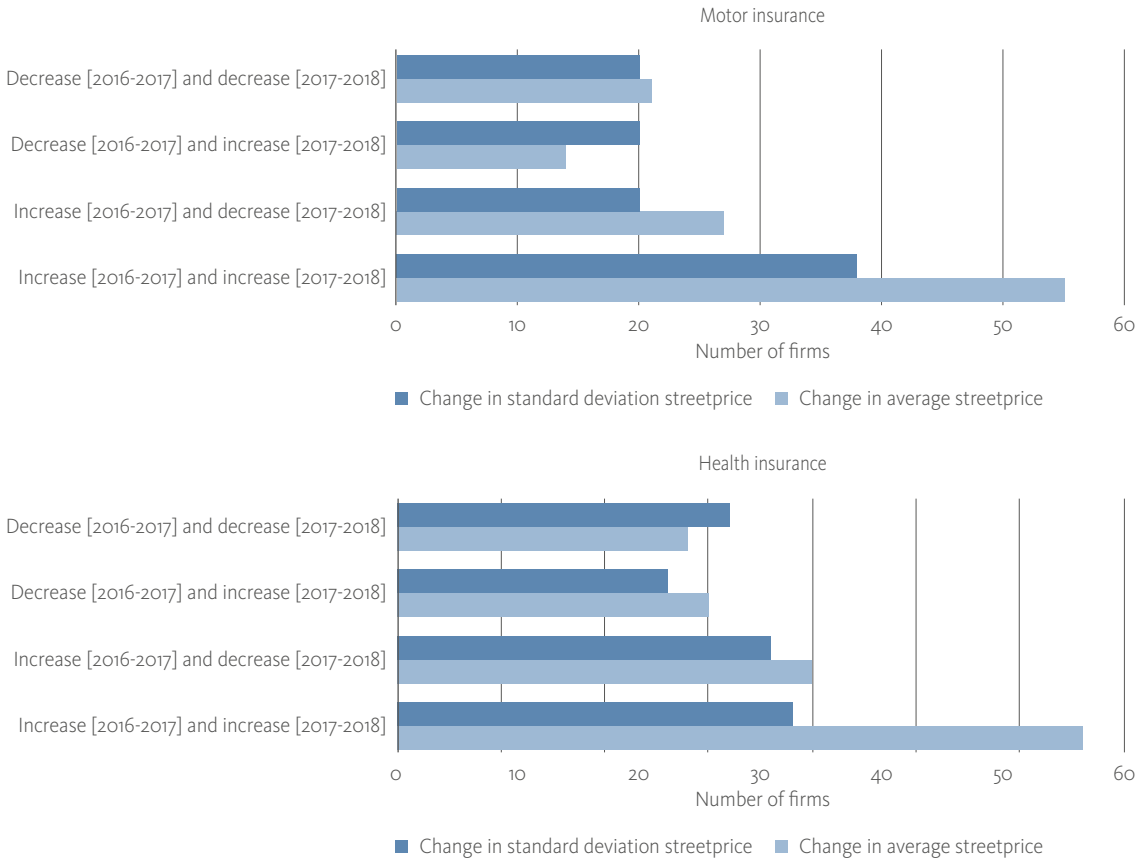
As far as the quality of the quantitative data provided by firms is concerned, many firms had to be left out of the analysis due to data quality issues like missing data-point or misinterpretation of the questions. For some parts of the quantitative question the data of about 70 to 120 firms could be used for motor insurance. For health insurance, the number of useful data points in combination with the quality were in the most cases too low for a proper analysis. Most of the analysis below is therefore based on the motor insurance survey.

Overall, the outcomes are very sensitive to the quality of the reported data, interpretation of the question by the firms, definitions of the different terms and the diversity of the portfolio (motor and health). Therefore it is important to treat this information very cautiously and do not adopt conclusions solely based on this information. Indeed, even with the use of a smaller portfolio or with representative persons, it will still be difficult to proof that a change in standard deviation is solely driven by BDA.

Evolution of average premiums and standard deviation

The standard deviation of the average premium measures the spread between the lower and the higher premiums (i.e. lower and higher risks pools, or lower and higher risk consumers). The hypothesis is that the standard deviation will increase overtime because of BDA. EIOPA has collected from insurance firms the data on standard deviation in motor and health insurance lines of business for the years 2016, 2017, and 2018. The results are shown in the graphics below.

Figure 35 – Evolution of the average premium and standard deviation



Source: EIOPA BDA thematic review

There are more firms that do not show an increase in the standard deviation in both consecutive years than firms that show an increase in both years. However, in motor insurance the group of firms that experience an increase in the standard deviation in both consecutive years is quite significant. Based on this information, and taking into account that other factors such as inflation or competition dynamics can also affect the evolution of the standard deviation, it is not possible to conclude that an increase in standard deviation is driven by BDA.

Number of rating-factors, risk-pools and change in standard-deviation in motor insurance

The hypothesis is that an increase in the number of risks-pools and rating factors leads to an increase in the standard deviation. The following analysis is done by combining the quantitative information on standard deviation and the qualitative information provided by firms about the evolution on the number of risk-pools and rating in the last three years.

One could argue that an increase in the number of rating-factors would also cause an increase in the number of risk-pools. Therefore, it was expected to see that an increase in the number of rating factors should have at least an increase in the number of risk-pools. For example, if a firm has only 3 yes/no rating factors, the number of risk-pools is 2x2x2=8, if the number of rating factors increases to 4 the number of risk-pools doubles to 2x2x2x2=16.

In the table below there is an overview of the number of companies and their answers. Some firms reported an increase in the numbers of rating factors but then reported 'no-change' in the number of risk-pools (nine firms) or even a reduction in the number of risk-pools (three firms). This outcome is not in line with what we expected to see. Therefore it is important to draw any conclusions from the data.

Figure 36 – Relationship between risk-pools and rating factors

Change in risk-pools	Change in rating-factors		
	Less	**No-change**	**More**
Less	2	1	3
No-change	1	21	9
More	1	11	32

Source: EIOPA BDA thematic review

The change in number of risk-pools with the change in standard deviation was also compared. Theoretically, an increase in the number of risk-pools will increase the standard deviation of the 'street price' of new motor insurance policies. The potential changes were clustered into three broad categories: decrease, no-change and increase in the number risk-pools or rating factors. This analysis is shown in Figure 37.

Figure 37 – Relationship between number of risk pools and standard deviation

Change in number of risk pools	Number of firms	Average change	Lowest value in the data	Highest value in the data	Standard deviation
Decrease	5	1.1%	-26.4%	35.9%	23.2%
Increase	46	6.4%	-56.5%	68.9%	19.9%
No change	40	1.8%	-100.0%	68.8%	23.1%

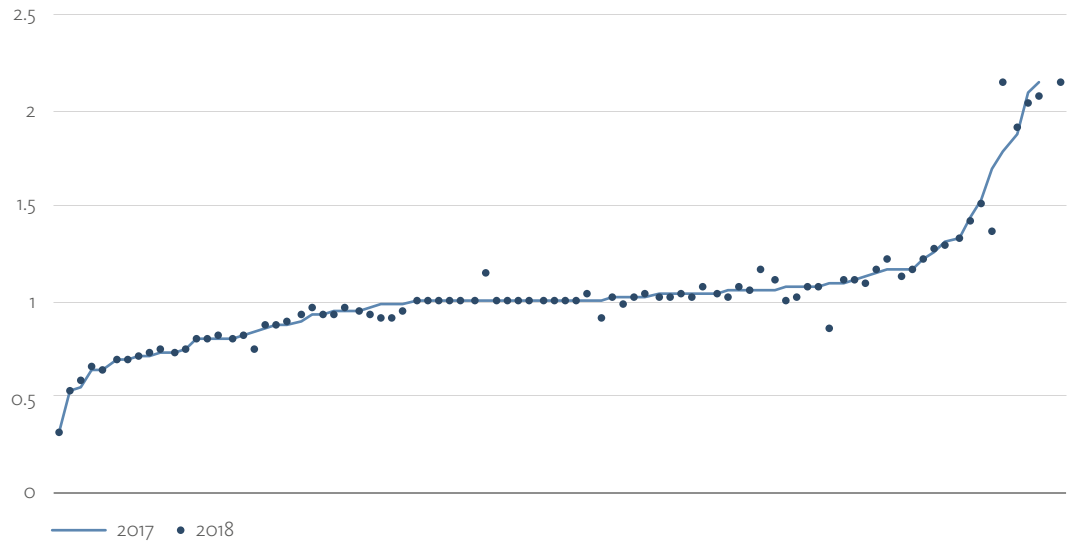
Source: EIOPA BDA thematic review

In total 46 firms reported an increase in the number of risk-pools and the average change in standard deviation was 6.4%. Therefore, it could potentially stated that an increase in the number of risk-pools leads to an increase in the standard deviation (and vice-versa). However, the data is very diverse and shows a high variability. Also the average change is highly dependent of an individual high or low annual growth value. Furthermore, when the same exercise was done with the 'technical price' instead of the 'street price', the results were very different.

The relationship between the technical and street price for motor insurance

Theoretically, it could be assumed that the street price would be higher than the technical price, otherwise a firm sells policies with an expected loss. The reported data showed another picture. In the figure below we plotted a frequency distributions of 85 firms, comparing the street- and technical price (ratio) of the new policies sold in 2017 (blue line) with the same ratio in 2018 (30/6 and black dot).

Figure 38 – Technical price vs street price ratio



Source: EIOPA BDA thematic review

Figure 38 shows that there are many insurance firms with a (much) higher technical price than the street price. There were 34 firms with a ratio below the 100% [street price exceeds the technical price], 14 firms which do not have a difference and have the same price for both, and 38 firms with a ratio above the 100% [street price is below the technical price].

The change in ratio from year to year does not point to one direction or shows a trend (see table below), 39 firms show a decrease in the ratio and 30 firm show an increase in the ratio. Statistically it is not possible to prove anything when the data does not show a particular direction or pattern. The same analysis based on the street- and technical prices of total policies (i.e. not only new policies) showed the same weak results.

ANNEX III: BDA REGULATORY FRAMEWORK

Cross-sectoral legislation

The Joint ESAs report on the use of Big Data by financial institutions provides a comprehensive overview of the cross-sectorial legislative framework applicable in the area of BDA.[35] This includes the following:

- General Data Protection Regulation (GDPR),
- Network and Information Systems Directive (NIS),[36]
- Directive on Distance Marketing of Financial Services,[37]
- Regulation on a framework for the free flow of non-personal data in the European Union,[38]
- the upcoming E-Privacy Directive,[39] or
- the Unfair Commercial Practices Directive (UCPD).[40]

The Council Directive 2004/113/EC on equal treatment between men and women in access to and supply of goods and services[41] and the recently approved Regulation on a framework for the free flow of non-personal data in the European Union[42] are also relevant in this context.

It is not the purpose of the present thematic review to repeat the regulatory analysis already included in the Joint ESAs report; therefore this section focuses on analysing some of the key provisions of the GDPR that insurance firms mentioned when they were asked how would the GDPR help them address some of the challenges arising from BDA. Indeed, several insurance firms believe that data quality/accuracy, fairness as well as ethical considerations were already important under the previous privacy legislation, but the GDPR has further enabled them to increase their data governance processes and transparency vis-à-vis consumers.

In this regard, several insurance firms referred to the principles regarding the processing of personal data included on Article 5(1) GDPR. This article includes very important principles such as the principle of purpose limitation, requiring firms to have and inform individuals about the specific purposes for processing the data and ensuring that any further processing is compatible with the original purpose. Firms must also only collect and process the personal data that is necessary to fulfil that purpose (principle of data minimisation).[43]

Article 5(1) GDPR also comprises the principle of integrity and confidentiality: firms must install appropriate technical and organisational safeguards that ensure the security of the personal data. Furthermore, firms must also ensure the personal data is stored for no longer than necessary for the purposes for which it was collected (principle of storage limitation) and that the personal data used is accurate and up-to-date (principle of data accuracy).

Firms must always treat consumers fairly and transparently when processing their data (principle of fairness and transparency), and must be responsible for and be able to demonstrate compliance with the above-mentioned principles (principle of accountability). Moreover, firms must have in place the adequate governance measures that ensure the protection of consumer's privacy right from

35 Report on the use of Big Data by financial institutions, Joint Committee of the ESAs, 15 March 2018, https://eiopa.europa.eu/Publications/Other%20Documents/JC-2018-04%20Joint%20Committee%20Final%20Report%20on%20Big%20Data.pdf

36 Directive 2016/1148 on Security of Network and Information Systems

37 Directive 2002/65/EC on distance marketing of consumer financial services.

38 European Commission press release: EU negotiators reach a political agreement on free flow of non-personal data, 28 June 2018, http://europa.eu/rapid/press-release_IP-18-4227_en.htm

39 Directive (EC) 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications)

40 Directive 2005/29/EC of 11 May 2005 concerning unfair business-to-consumer commercial practices.

41 Council Directive 2004/113/EC on equal treatment between men and women in access to and supply of goods and services, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32004L0113. It is important to note that on 1 March 2011, the Court of Justice of the European Union declared invalid as from 21 December 2012 an exemption in EU equal treatment legislation which allowed Member States to maintain differentiations between men and women in individuals' premiums and benefits.

42 European Commission press release: EU negotiators reach a political agreement on free flow of non-personal data, 28 June 2018, https://europa.eu/!HC67XW

43 See the recent Decision from the Bundeskartellamt in Germany, which has imposed restrictions in the processing of user data on Facebook: https://www.bundeskartellamt.de/SharedDocs/Meldung/EN/Pressemitteilungen/2019/07_02_2019_Facebook.html?nn=3591568

the start of the processing (principle of data protection by design and by default).

As far as the lawfulness of the processing of personal data is concerned, insurance firms explained that consumers commonly provide their consent (Article 6(1)(a) GDPR) to the processing of their personal data when they accept the terms and conditions of their insurance policy and/or via dedicated privacy notices. However, some firms noted that, other than obtaining consumer's consent, they typically make use of most of the six legal grounds recognised under Article 6 GDPR. This would be the case for instance in order to comply with legal obligations, such as the ones steaming from Solvency II, or to pay claims during the duration of the insurance contract. Insurance firms may also process personal data for legitimate purposes, for instance in the case of fraud prevention.

Insurance firms considered that GDPR has introduced relevant requirements concerning the transparency in the processing of personal data (e.g. Articles 13 and 14 GDPR).[44] Firms explained how consumers need to timely, appropriately and transparently be informed, from their first touch-point with the firm, about how their personal data is processed, including both data from internal and external data sources. This also included the requirement of informing them about the existence of automated decision making processes (i.e. BDA), as well as providing them with meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing.

Furthermore, firms explained that the GDPR has recognised consumers a number of rights relevant in the context of BDA, including the right of access to their data, in order to verify the accuracy of the data and the lawfulness of the processing (Article 15 GDPR). Consumers can also request modifications (Article 16 GDPR) or even to object to auto-mated decision making processing in certain circumstances (Article 18 GDPR). According to Article 22 of GDPR, consumers will also be able to (i) ask financial institutions that a human intervene in the profiling, to (ii) express their point of view and (iii) contest a decision based on profiling.

Moreover, several insurance undertakings highlighted the importance of the requirement to appoint a dedicated Data Protection Officer (Article 37 GDPR) and to develop a data protection impact assessment (DPIA) (Article 35 GDPR); the DPIA is the process that evaluates the impact on rights and freedoms of individuals before launching any new data processing activities. DPIAs examine considerations relating to the accuracy of the data, both in terms of new data being received, how that data is matched to any existing customer records and how it can be assured the outcome is also accurate.[45]

It is also important to mention that the Article 9 GDPR introduces some restrictions concerning the processing of personal categories of data. These restrictions are in line with Article 21(1) of the European Union Charter of Fundamental Rights, which establishes that 'any discrimination based on any ground such as sex, race, colour, ethnic or social origin, genetic features, language, religion or belief, political or any other opinion, membership of a national minority, property, birth, disability, age or sexual orientation shall be prohibited.' According to Article 21(2), discrimination on grounds of nationality shall also be prohibited.

Last but not least, Article 5(2) of the Unfair Commercial Practices Directive defines unfair commercial practices as those that are contrary to the requirements of professional diligence and are likely to distort the economic behaviour of an average consumer. Furthermore, Article 5 (3) of the same Directive specifies that special consideration should be given to vulnerable consumers.

Insurance legislation

At international level, the Insurance Core Principles (ICP) developed by the International Association of Insurance Supervisors (IAIS) contain relevant provisions. In particular, ICP 19 states that insurance supervisors must ensure that 'insurance firms and intermediaries, in their conduct of insurance business, treat customers fairly, both before a contract is entered into and through to the point at which all obligations under a contract have been satisfied.'

At European level, Article 41 of the Solvency II Directive requires from all insurance and reinsurance companies 'to have in place an effective system of governance which provides for sound and prudent management of the business.' Articles 38 and 49 Solvency II Directive also sets out the requirements regarding the outsourcing of functions and activities (e.g. collaboration with data vendors).

44 These provisions are complemented by the guidelines developed by the European Data Protection Board on transparency (https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=622227) and consent (https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=623051)

45 Article 29 Working has provided further guidance on how to develop a DPIA, which does not necessarily include an anti-discrimination nor compliance with the principle of fair treatment assessment: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=611236)

Furthermore, the Article 19 of the Delegated Regulation (EU) 2015/35[46] containing implementing rules for Solvency II establishes detailed data accuracy/quality requirements, but only in relation to the data used in the calculations of the technical provisions.

The product oversight and governance (POG) requirements under the Insurance Distribution Directive (IDD) are also relevant (e.g. identification of the target market). Finally, it is also important to mention Article 17(1) of the IDD requiring from insurance distributors to act honestly, fairly and professionally in accordance with the best interests of their customers.

Relevant provisions may also be found in national insurance legislation; for example in France, the Evin Law specifies the data that can and cannot be used by health insurance undertakings. Moreover, Belgium passed a law in April 2014 according to which insurance undertakings need to publish the segmentation criteria for 6 types of insurance contracts.

46 Delegated Regulation (EU) 2015/35, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L:2015:012:TOC

ANNEX IV: HOW BIG TECH FIRMS COULD ENTER THE INSURANCE MARKET

One of the questions in the industry questionnaire asked insurance firms to which extent they see Big Tech firms entering the insurance market in the next three years and how could this impact their BDA processes. The majority of the respondents stated that they have observed an increasing interest of some Big Techs in entering the insurance market, although there were diverging views on whether they would eventually enter it, in what form and when.

Respondents acknowledge that Big Tech Firms have an unusually large customer base and access to large amounts of different types of data which are not available to traditional insurance companies; they can leverage on data collected from several sources and different types of markets. Shall Big Techs decide to enter the insurance market, many insurance firms consider that this would take place in the form of intermediaries / brokers / price comparison websites; they consider that they could disrupt the distribution of insurance products by selling insurance products through their platforms.[47]

In this scenario, Big Techs would obtain an insurance distribution license and then collaborate with insurance undertakings, which would focus on risk-taking regulated activities. Some firms consider that this could eventually lead to reduced underwriting margins for insurance firms. They argue that large platforms with strong bargaining power could potentially engage in orchestration and gatekeeping practices; e.g. defining the 'rules of the game' by favouring certain products in the ranking criteria of their platforms or by controlling the entities that can sell products through their platforms.

Moreover, if Big Tech firms would become increasingly active in the distribution of insurance policies, some insurance firms believe that they would end up themselves having less contact with consumers and therefore less access to key consumer behavioural data. This would therefore affect their own BDA processes, since this information is commonly used for supporting insurance firm's operations in different processes across the insurance value chain.

Another group of respondents consider that Big Tech firms could try to leverage on the extensive data they already have about consumer's behaviour (e.g. shopping habits, web searches, health data etc.) in order to predict future claims and therefore enter the market with an insurance undertaking license. More particularly, some respondents believe that they are more likely to do this firstly in the health insurance sector. Other insurance firms consider that Big Tech firms could focus on niche and special purpose products, such as IoT or Blockchain-based products.

On the contrary, some of the respondents stated they hardly expect large Tech Firms to enter the insurance market. There is the feeling that Big Tech Firms generally tend to avoid highly regulated markets because the required level of transparency towards the regulators is too high. Furthermore, the high-level fragmentation and complexity of European and national insurance regulatory framework could be seen as another major obstacle for them. In addition, their lack of historical claims data could also deter them from entering the insurance market, although some respondents argued that they could solve this gap via mergers and acquisitions of established insurance firms.

Interestingly, one insurance firm notes that Big Tech Firms will not enter the insurance market because their entire business is 'data' and not 'insurance'. They argue that although Big Tech firms have access to relevant datasets and state-of-the-art technological expertise, they do not have the insurance knowledge. Therefore, some insurance firms consider that Big Techs will rather focus in providing data and technological solutions (e.g. cloud computing, AI/ML technology, advertising) to insurance firms on a B2B basis.

47 The Metcalfe's law or network effect describe the effect that the value of a network increases proportionally to the square of its numbers of users. Furthermore, a data driven business model can be enhanced when more data is available. Both effects can lead to strong trend to monopolies in platform economics

ANNEX V: GLOSSARY

The following terms were used for the purpose of the present thematic review, taking into account the work of international institutions such as the Bank of International Settlements (BIS). EIOPA acknowledges that there is not a unique definition of these terms and/or that they are constantly evolving. Therefore EIOPA might use a different terminology in further work on this area.

Artificial intelligence	IT systems that perform functions usually performed by human capabilities. AI can ask questions, discover and test hypotheses, and make decisions automatically based on advanced analytics operating on extensive data sets. Machine learning (see below) is one subcategory of AI.*
Artificial Neural Networks (ANNs)	Artificial Neural Networks (ANNs) are a type of machine learning algorithms that are commonly represented in a similar fashion as the human brain; they operate with an input layer, one or more unknown hidden layers, and an output layer. In what is known as a "feedforward network", the information flows from the input layer, through the hidden layer into the output layer (for further information please refer to point 2.2.1)
Big Data Analytics (BDA)	Large volumes of data that can be generated, processed and increasingly used by digital tools and information systems for making predictive, descriptive and prescriptive analysis. This capability is driven by the increased availability of structured data, the ability to process unstructured data, increased data storage capabilities and advances in computing power.
Descriptive analytics	The use of data aggregation and data mining to provide insights into the past and answer what has happened.
Health insurance	"(29) Health insurance obligations where the underlying business is pursued on a similar technical basis to that of life insurance, other than annuities stemming from non-life insurance contracts and relating to health insurance obligations" **
Internet of Things (IoT)	Is the networking of telematics devices, vehicles, buildings, and other items embedded with electronics, software, sensors, wearables actuators, and network connectivity that enable these objects to (a) collect and exchange data and (b) send, receive, and execute commands
IoT-based insurance products	Insurance products based on IoT sensor devices to measure consumer's behaviour and environment to perform risk assessments and price discount rewards. For instance, this would be the case of Pay-As-You-Drive (PAYD) and Pay-How-You-Drive (PHYD) products in motor insurance, or Pay-As-You-Live (PAYL) products in health insurance.
Machine learning	Machine learning (ML) is the ability of computers to learn from data through appropriate algorithms. This allows them to build a model of their world and better solve their intended tasks. Approaches of ML can be characterized by the dimensions of the task (differentiating fundamentally between classification, regression and clustering), the data types (special approaches exist for example for text, language and image data) and the algorithms (how is the problem solved technically). †
Motor insurance	"(4) Insurance obligations which cover all liabilities arising out of the use of motor vehicles operating on land (including carrier's liability) and insurance obligations which cover all damage to or loss of land vehicles (including railway rolling stock).†
New policies in period	New insurance policies concluded in a given period, without taking into account renewed policies.
Non-personal data	Any kind of data that is not personal data

Personal data	Personal data means any information relating to an identified or identifiable natural person; an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person[††]
Predictive analysis	Making future predictions by studying recent and historical data.
Prescriptive analytics	The use of data aggregation and data mining to provide recommendations of one or more courses of action and showing the likely outcome of each decision.
Price optimisation	Adjustments to the technical price to create the street price using factors which are unrelated to the risk of loss (including the add on of discretionary costs such as fees, charges and commissions).
Pricing or pricing practices	The method and setting of the price. This includes setting the technical price and making any optimisation adjustments including the add on of discretionary costs such as fees, charges and commissions to determine the street price
Rating factor	Any factor that is involved in the process of pricing of an insurance policy, and influences the premium paid by the consumer.
Robo-advisors	Phenomenon whereby advice is provided to consumers without, or with little, human intervention and providers rely instead on computer-based algorithms and/or decision trees.
Street price	The actual or ultimate price paid by the consumer.
Technical price	Pricing using actuarial rating factors, such as expected claims costs, commissions, profit load and cost of capital.
Total policies at the end of period	Total number of policies at the end of a given period, including both new policies and renewed policies
Virtual assistant / Chatbot	A computer program that simulates human conversation through voice commands or text chats or both. Chatbots are typically embedded into messaging applications.

* Based on the definition used in the Bank of International Settlements report "Implications of fintech developments for banks and bank supervisors", August 2017, https://www.bis.org/bcbs/publ/d415.htm

*[*] Commission Delegated Regulation (EU) 2015/35 of 10 October 2014 supplementing Directive 2009/138/EC of the European Parliament and of the Council on the taking-up and pursuit of the business of Insurance and Reinsurance (Solvency II), pages 227 and 228, http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L:2015:012:FULL&from=EN

[†] Delegated Regulation (EU) 2015/35, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L:2015:012:TOC

[††] Article 4(1) GDPR

* Commission Delegated Regulation (EU) 2015/35 of 10 October 2014 supplementing Directive 2009/138/EC of the European Parliament and of the Council on the taking-up and pursuit of the business of Insurance and Reinsurance (Solvency II), pages 227 and 228, http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L:2015:012:FULL&from=EN

ANNEX VI: LIST OF ACRONYMS

AI	Artificial Intelligence
ANNs	Artificial Neural Networks
BDA	Big Data Analytics
BIS	Bank of International Settlements
CRM	Customer Relationship Management systems
CUI	Customer User Interfaces
DPIA	Data Protection Impact Assessments
DPO	Data Protection Officer
EBA	European Banking Authority
EEA	European Economic Area
EIOPA	European Insurance and Occupational Pensions Authority
ESAs	European Supervisory Authorities (EBA, ESMA and EIOPA)
ESMA	European Securities and Markets Authority
FNOL	First Notice of Loss
GDPR	General Data Protection Regulation
GLM	Generalised Linear Models
GWP	Gross Written Premiums
IaaS	Infrastructure as a Service
IAIS	International Association of Insurance Supervisors

ICP	Insurance Core Principles
IDD	Insurance Distribution Directive
IoT	Internet of Things
IVR	Interactive Voice Response
JC	Joint Committee of the ESAs
ML	Machine Learning
MTPL	motor third-party liability insurance
NAIC	National Association of Insurance Commissioners
NCA	National Competent Authority
NLP	Natural Language Processing
OBD	On Board Device
OCR	Optical Character Recognition
PaaS	Platform as a Service
PAYD	Pay-As-You-Drive
PAYL	Pay-As-You-Live
PHYD	Pay-How-You Drive
RPA	Robotic Process Automation
SaaS	Software as a Service
UBI	Usage-based insurance

GETTING IN TOUCH WITH THE EU

In person

All over the European Union there are hundreds of Europe Direct information centres. You can find the address of the centre nearest you at: https://europa.eu/european-union/contact_en

On the phone or by email

Europe Direct is a service that answers your questions about the European Union. You can contact this service:

– by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),

– at the following standard number: +32 22999696, or

– by email via: https://europa.eu/european-union/contact_en

FINDING INFORMATION ABOUT THE EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website at: https://europa.eu/european-union/index_en

EU publications

You can download or order free and priced EU publications from: https://publications.europa.eu/en/publications. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see https://europa.eu/european-union/contact_en).

EU law and related documents

For access to legal information from the EU, including all EU law since 1952 in all the official language versions, go to EUR-Lex at: http://eur-lex.europa.eu

Open data from the EU

The EU Open Data Portal (http://data.europa.eu/euodp/en) provides access to datasets from the EU. Data can be downloaded and reused for free, for both commercial and non-commercial purposes.

EI-02-19-220-EN-C

EUROPEAN INSURANCE AND OCCUPATIONAL PENSIONS AUTHORITY

Westhafenplatz 1,
60327 Frankfurt am Main, Germany

Publications Office
of the European Union