A quantitative analysis on the effect of COVID-19 in a private health insurance plan expenditure

Davide Biancalana[1] **· Fabio Baione**[1]

Abstract

This paper explores the effect of COVID-19 on health care expenditure using data from a private Health Insurance Plan (HIP). As well known, at the beginning of the COVID-19 pandemic, governments had to rely on Non-Pharmaceutical Interventions against the spread of the virus. However, the stringency of lockdowns differed across space and time as governments had to adjust their strategy dynamically to the country-specific development of the crisis. These strategies have strongly changed the policyholders' behavior; however, after this period, a fundamental question is whether the policyholder behavior will return to a status quo (i.e. in traditional care delivery). We analyze these effects using a "pre-post" quantitative study using longitudinal data collected from 2017 to 2021. We consider as a consumption measure the health care expenditure amount within several types of health services, coming from a group of insured persons, followed overtime every quarter, and separating the effect per gender and age. Moving in this direction, the purpose of our contribution is to investigate if the traditional actuarial approach for assessing the loss cost, based on the Generalized Linear Models, could predict the effect on the health care expenditure due to COVID-19 and the capacity to which a HIP can anticipate these uncertainties. Our results provide a comprehensive picture of the different effects of COVID-19 on the health services offered by the HIP, as well as on the behavior ofpolicyholders during and after the pandemic period.

Keywords COVID-19 · Health expenditure · Generalized linear models · Two part model · Tweedie

1 Introduction

The global spread of COVID-19 developed into the most serious threat to health since World War II. This paper deals with the assessment of health expenditure changes due to COVID-19 in a private Health Insurance Plan (henceforth HIP) and to understand if

✉ Davide Biancalana
davide.biancalana@uniroma1.it

1   Department of Statistics, Sapienza University of Rome: Universita degli Studi di Roma La Sapienza, Viale Regina Elena 295, Rome, Italy

the policyholder behavior has been modified after the pandemic period; other researches have treated about similar issue.

In Banthin et al. (2020) a study was carried out on the effects of employment losses on health insurance coverage in the USA. The authors have calculated that 48 million people will live in families with a worker who experiences a COVID-19-related job loss in the last three quarters of 2020. Of them, 10.1 million lose employer coverage tied to that job. They have estimated that 32 percent of these people switch to another source of employer coverage through a family member.

In Hansen et al. (2021) potential inter-dependencies between the public health care system and the efficiency of lockdown measures are investigated. The results indicate that there is a trade-off between the stringency of the lockdown and the prevailing health expenditures. Less stringent lockdowns in countries with higher health expenditures had a similar impact on mortality as more stringent lockdowns in countries with lower health expenditures. They show that the effects of lockdown interventions were insignificant in developing countries with per capita health expenditure below the mean.

In Bundorf et al. (221) is assessed the extent to which insurance coverage changed between mid-April and December 2020 in the USA. Furthermore, a study was carried out about differences by prepandemic family income, age and ethnicity based on evidence of the pandemic's disproportionate labor market effects across groups. The results indicate that much of the overall decline in coverage took place within a short 3-month period early in the pandemic. While employer-sponsored coverage declined throughout 2020, the decline was more fully offset by increases in other sources later in the year.

In Zuo and Zhai (2021), the work investigates the impact of China's COVID-19 treatment policy on the sustainability of its Social Health Insurance (SHI), explores influences of the policy on Wuhan's system, and discusses the effects of an assumed equivalent emergency on SHI funds for five other provincial capital cities in China. The results suggest the integration of insurance schemes and provincial pooling, fund balance adjusting and an emergency safety net is also advised. This paper contributes to the literature from a different point of view by focusing on the pre and post effects of COVID-19 on health care expenditure using data from an Italian private HIP. Our first quantitative question is: "How does COVID-19 impact the policyholder behavior of private health services after lockdown measures?"

The answer to this question is based on the measurement of the health care expenditure analyzing the number of episodes (henceforth frequency component) and the payment for the occurred episodes (henceforth severity component). The analysis has been carried out by observing the frequency and severity trends of the same insured persons during the pre-pandemic period 2017–2019 and mid/post-pandemic period 2020–2021 for several medical-care services.

Moreover, it is important to note that one standard method in actuarial literature is modelling the insurance expenditure by a two-part model (Frees 2009; Frees et al. 2013, 2016). Two-part models are useful for modelling the expectation of the interested expenditure variables through the product of the count estimates (used in episodes frequency) and the severity (used in episodes severity) estimates. As an alternative, to provide an estimate of the loss cost, also known as 'pure premium' using a single model it is possible to assume a Tweedie distribution. This distribution will help us in modeling pure premium directly without any need for two different models. Tweedie distribution is a special case of exponential dispersion models and is often used as a distribution for GLMs. It can have a cluster of data items at zero and this particular property makes it useful for modeling claims in the insurance industry. This model can also be applied in

other use cases across industries where you find a mixture of zeros and non-negative continuous data points.

Hence, our second research question is: "Could a standard actuarial model, based on a Tweedie GLM, predict the effects of COVID-19 on the estimated expected loss costs?"

To this aim we conduct a quantitative analysis fitting a Tweedie GLM to our pre-pandemic data and then we forecast the health care expenditure in the pandemic and post-pandemic period; the model used allows for point estimate and confidence interval estimate of the interested loss variables. Furthermore, this analysis will be carried out considering not only calendar year and quarter but also policyholder features such as gender and age to asses which gender-age risk profiles have been most affected by pandemic outbreak. The rest of the paper is organized as follows: Sect. 2 introduces data and methods, Sect. 3 illustrates the results of a numerical application on a real insurance database and Sect. 4 concludes the work.

2 Data and Methods

2.1 Data

Health care expenditure over the period 2017–2021 is taken from a private HIP operating in Italy in a specific North region. The longitudinal data track the same sample of 16, 206 policyholders whose health care expenditures are collected quarterly. Longitudinal data allow for the measurement of within-sample change over time.

Italian's HIPs usually offer supplementary and complementary health care protection to the services provided by the National Health Service (NHS); the latter assures free health care for its citizens. The health care services covered by the plan we analyzed are grouped as follows:

These classes contain various medical-care episodes that can vary in the type of services and terms of frequency and severity. Then, medical-care episodes in the same class of service (henceforth class or group) are considered homogeneous. As shown in Table 1 the Specialist Medical Visits (SMV) represents the highest frequency group, while Rehabilitation & Physical Care (RBC) are the highest severity group.

The input data consist of the number (counts) and cost (expenditures) of invoices per quarter period and policyholder, distinguished by the group of services provided. The data set contains 66.69% of females and 33.31% of males, whereas the age distribution is quite similar, as shown in the following Fig. 1:

It is worth noting that during the COVID-19 pandemic, Italy observed a hard lockdown period (Phase 1) between 9 March and 3 May 2020. Indeed, on 11 March, a Prime

Table 1 Group of services (in alphabetic order) and annual number and amount of expenditure

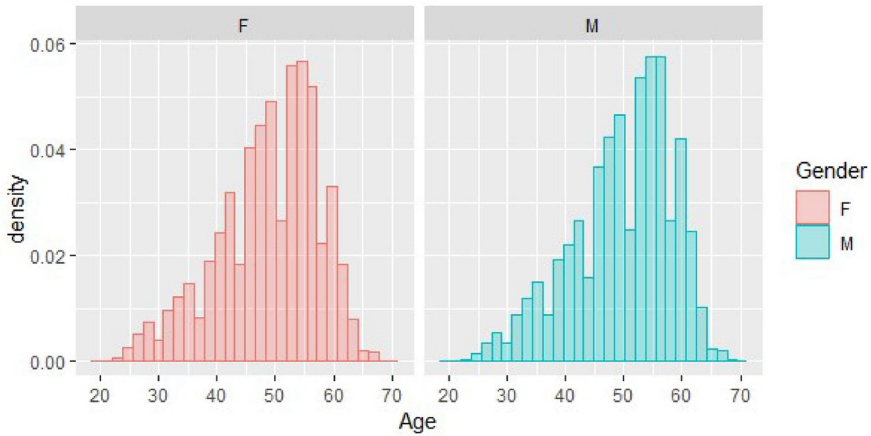| Group of services	Acronym	Number	Expenditure
Dental care	DC	2,832.40	250,470.55
Oral hygiene	OH	1,233.40	201,313.33
Rehabilitation & physical care	RPC	4,075.00	1,817,534.79
Supplementary care	SC	1,262.00	131,322.72
Specialist medical Visits	SMV	13,707.20	873,586.79

Fig. 1 Gender and age distribution

Ministerial Decree was published, nicknamed the "I Stay at Home Decree," which provides for the suspension of everyday retail, commercial activities, educational activities, and catering services and prohibits gatherings of people in public places or places open to the public. The subsequent Phase 2 relaxed the containment measures (May 4–June 14), with a gradual relaxation of the previous containment measures, as the epidemic curve is in a downward phase. Phase 3 (June 15–October 7) and the subsequent phases consist of coexistence with COVID-19 that still loosens the containment measures, leading to a degree of "normality" in daily activities.

Therefore, according to the set theory, we denote A as the set of quarters before the pandemic (i.e., 2017–2019). Hence, A contains the "Pre" pandemic data. Then, we denote B and C as the quarters during (i.e., 2020) and after (i.e., 2021) the pandemic, respectively. Therefore, B and C contain the "During" and "Post" pandemic data. We analyze the behavior of the standard set ($A \cup B \cup C$) over a group of services.

2.2 Methodology: the frequency-severity model and the tweedie regression

Many insurance datasets are characterized by information about how often episodes arise in addition to the corresponding size. A standard model in actuarial literature to analyze such data consists of the so-called frequency-severity (or two-part) model, which separately models: the number of episodes or frequency per unit of exposure (i.e., time, value) and the per episode amount or severity. The frequency-severity method uses historical data to estimate the expected number of episodes and the expected cost of each episode during a given period.

In our context, following the collective risk theory approach (see, for instance, Daykin et al. (1993)), given a generic group of services k and n policyholders. Let i index the generic policyholder ($1 \leq i \leq n$).

The ith policyholder expenditure (or aggregate loss) can be seen as the sum of the single expenditures, $_kY_{ij}$, for the single service requested j in the group k, as follows:

$$_kS_i = \sum_{j=1}^{_kN_i} {}_kY_{ij}, \tag{2.1}$$

where $_kN_i$ represents the random variable (r.v.) number of episodes for group of services k requested by the ith policyholder in a given period. The approach we use to investigate the effect of the pandemic is to analyze the distribution of the number and cost of services (as reported in formula (2.1)) of the single policyholder during a different set of periods (A, B, and C) previously defined. It is worth noting that, if there are no requests with policyholder i and service k, $_kN_i = 0$ implies that $_kS_i = 0$.

Assuming $_kY_{ij} \sim_k Y_i$ iid, respect to j, and independent of $_kN_i$ we have:

$$E(_kS_i) = E(_kN_i)E(_kY_i), \quad Var(_kS_i) = E(_kN_i)Var(_kY_i) + Var(_kN_i)E^2(_kY_i). \tag{2.2}$$

Finally, ignoring, for the sake of simplicity, the superscript k and subscript i. Suppose that $N \sim Poisson(\lambda)$, and each $Y \sim Gamma(\alpha, \theta)$ with α and θ shape and scale parameters, respectively. By using iterated expectations, Eq. (2.2) can be expressed as:

$$E(S) = \lambda\alpha\theta \quad \text{and} \quad Var(S) = \lambda\alpha(1 + \alpha)\theta^2. \tag{2.3}$$

Now, define three parameters μ, ϕ, and p through the following reparametrization

$$\lambda = \frac{\mu^{2-p}}{\phi(2-p)}, \quad \alpha = \frac{2-p}{p-1}, \quad \theta = \phi(p-1)\mu^{p-1}, \tag{2.4}$$

it is easy to show that $E(S) = \mu$ and $Var(S) = \phi\mu^p$ (see Kaas 2005; Tweedie 1956). As a consequence, $S \sim Tw(\mu, \phi)$ denotes a Tweedie random variable with mean $\mu > 0$, variance $\phi\mu^p$ ($\phi > 0$), where ϕ is the so-called "dispersion parameter", and $p \in (-\infty, 0] \cup [1, \infty)$ the power parameter. Its distribution belongs to the exponential family, allowing us to use the Tweedie distribution with GLMs to model $E(S_i)$ and $Var(S_i)$ for each ith risk profile. Tweedie distribution for two-part data is one of the most widely used mixture distributions in insurance claims modeling (Frees 2014). Indeed, this distribution will help us model the aggregate loss or pure premium directly without needing two different models for frequency and severity, respectively.

2.3 Model application

As previously stated, frequency-severity modeling is a standard actuarial approach to analyze insurance datasets. Anyway, the utilization of health-care services can be influenced by a set of explanatory variables categorized as demographic and geographic, among others. For example, demographic factor like age shows that, in some service, an increase in age results in an increasing impact on the demand for health care. Similarly, gender can be treated as a proxy for inherited health and different habits in maintaining health. We model this independent explanatory variables by the row vector $\mathbf{x}_i$.

Since analyzing health expenditure rating factors such as gender and age are fundamental, a univariate analysis for each risk class would require $2 \cdot (\omega - 1)$ two-part models, where $\omega - 1$ is the maximum age. Even in such a simple case, a univariate approach for each homogeneous risk class is not feasible; a multivariate approach is more appropriate. An industry-wide approach is based on multivariate regression models such as Generalized Linear Model (GLM), which has found extensive use in actuarial practice.

As the expenditure per single policyholder and group of services k has a mass probability at zero and a positive, continuous, and often right-skewed component for positive values, Tweedie regression is a valid candidate to predict such an expenditure.

As demonstrated in Kurz (2017), Tweedie distribution fits health care cost data very well and provides better fit, especially when the number of non-users is low and the correlation between users and non-users is high. The author states that a common way to counter this is the use of Two-part or Tobit models, which makes interpretation of the results more challenging. Tweedie distribution provides an interesting solution to many statistical problems in health economic analyses.

The Tweedie regression is a particular GLM based on the assumption that the dependent variable $_kS_i$ is Tweedie distributed and that its mean is related to a set of covariates through a linear predictor $\mathbf{x}_i$ with unknown coefficients $_k\boldsymbol{\beta}$ and a link function g. The conditional expected expenditure and variance for the ith policyholder and kth group of services are given by:

$$_k\mu_i = \mathrm{E}\left[_kS|\mathbf{x}_i\right] = \mathrm{E}\left[_kS_i\right] = g^{-1}(\mathbf{x}_i \cdot_k \boldsymbol{\beta}), \quad _k\sigma_i^2 = \mathrm{Var}\left[_kS|\mathbf{x}_i\right] = \mathrm{Var}\left[_kS_i\right] = \phi \cdot_k \mu_i^p \tag{2.5}$$

The estimate of the row-vector of regression parameters $_k\boldsymbol{\beta}$, of the dispersion parameter ϕ and the power p are provided by a Maximum Likelihood Estimation (MLE) approach (see Frees 2014).

In our particular case, the $\mathbf{x}_i$ contains information about **Gender "G"** (M, F), **Age "A"** (from 0 to 70), **Quarter "Q"** (Q1, Q2, Q3 & Q4) and **Year "Y"** (2017, 2018 & 2019) of the ith observation. For a given group of services $k = \{\mathrm{OH, SC, DC, RPC}\}$, we fit the following model:

$$\mathrm{E}\left[_kS|\mathbf{x}_i\right] = e^{\left(_k\beta_0 +_k\beta_{Y,w} \ln\left(Y_i\right) +_k\beta_G G_i + \sum_{w=1}^{\delta} {}_k\beta_{A,w}A_i^w +_k\beta_{Q2}Q2_i +_k\beta_{Q3}Q3_i +_k\beta_{Q4}Q4_i\right)} \tag{2.6}$$

Where as, for the group of services SMV we fit the following model:

$$\mathrm{E}\left[_{SMV}S|\mathbf{x}_i\right] = e^{\left(_{SMV}\beta_0 + \sum_{w=1}^{\delta} {}_{SMV}\beta_{Y,w}Y_i^w +_{SMV}\beta_G G_i + \sum_{w=1}^{\delta} {}_{SMV}\beta_{A,w}A_i^w +_{SMV}\beta_{Q2}Q2_i +_{SMV}\beta_{Q3}Q3_i +_{SMV}\beta_{Q4}Q4_i\right)} \tag{2.7}$$

where δ represents the degree of polynomial used for the numerical variables (i.e. Year, and Age). In this work, to avoid over parametrization, we set $\delta = 3$.

The model in Eq. (2.7) differs from the one in Eq. (2.6) just in the use of a polynomial function instead of a logarithmic function for the variable Year. The motivation for this choice is due to a better fitting of the model to the observed data.

The model coefficients are estimated using the MASS and tweedie packages of the R statistical software program (R Core Team 2016).

3 Results

3.1 Frequency results

To assess if COVID-19 has impacted the policyholder behavior in our HIP, we measured the frequency and severity components over the set of periods previously denoted as A, B, and C. Remembering that our longitudinal data refers to 16, 206 policyholders quarterly tracked, in Table 2 we show the quarterly average of the number of policyholders

Table 2 Data base basic information: frequency of zero episodes and severity quarterly mean over set

Set	Years	Description	DC	OH	SC	RPC	SMV
A	2017–2019	Policyholders w/o episodes	15,353	15,476	16,007	15,994	14,309
		No. of episodes	1,100	732	330	355	3,754
		Tot. expenditure	478,543.63	62,942.80	52,077.01	36,088.61	228,717.46
B	2020	Policyholders w/o episodes	15,470	15,595	16,047	16,059	14,719
		No. of episodes	858	613	251	238	2,748
		Tot. expenditure	404,173.08	56,714.85	41,948.61	25,586.73	185,565.63
C	2021	Policyholders w/o episodes	15,397	15,477	16,011	16,036	14,482
		No. of episodes	936	731	300	274	3,126
		Tot. expenditure	432,114.53	67,544.95	53,462.01	30,300.84	220,265.46

without any request of reimbursement over the selected set and group. Moreover, we show the quarterly average of the number of episodes and their expenditure.

At first sight, the analysis may suggest an unchanged policyholder behavior if we compare only the number of policyholders without episodes in sets A and C. Where as, the overall number of medical-care episodes in each group shows a general decrease. It is evident that the number of episodes in each class has declined during the COVID-19 pandemic, given the hard lockdown measures adopted by Italian governments during the first two quarters of 2020. It is less evident that in the post-pandemic period, the number of episodes slightly realigned to pre-pandemic data, especially for some groups. Indeed, by comparing 2021 values with pre-pandemic years, the number of episodes shows a significant reduction; for example, SMV and RPC reduce by 17% and 23%, respectively, although the amount of claims is not reduced by the same amount.

A step forward in the analysis consists of a graphical comparison of the numbers and average expenditure per person and period as proposed in Fig. 2 and 3. Each panel contains the chart of a group of services with the evidence of each set of periods examined.

The kernel density function of the frequency proposed in Fig. 2 empathize the differences among periods; indeed, set B and C show less "heavy-tailed" distributions compared to set A. These outcomes are confirmed by the descriptive statistics reported in Table 3.

The mean of set C seems to be realigned with pre-covid data, though high frequency groups, such as DC and SMV, are still lower than A values. This is probably due to the types of medical care episodes included in these classes that are not always strictly necessary and, therefore, can be postponed.

The mean of set C seems to be realigned with pre-covid data, though high-frequency groups, such as DC (0.06) and SMV (0.19), are still lower than A values (0.07 and 0.23, respectively). This is probably due to the types of medical care episodes included in these classes that are not always strictly necessary and, therefore, can be postponed.

Moreover, positive values for the skewness indicate data that are skewed right, which means that the right tail is long relative to the left tail. In addition, kurtosis over 3 indicates a "heavy-tailed" distribution. Significant skewness and kurtosis clearly indicate that data are not normal. Furthermore, A data figures out a heavy right tails per each group compared to the other sets with some exceptions. OH and SMV show distributions very similar
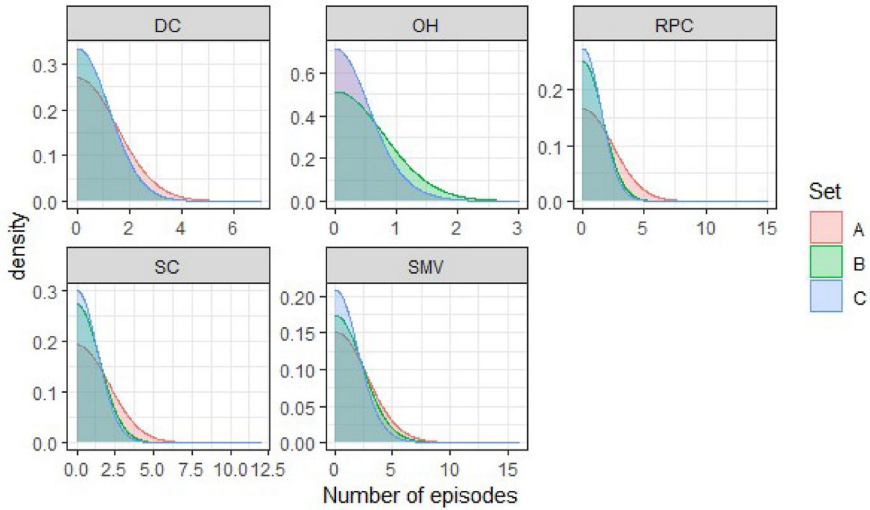
Fig. 2 Number of episodes distribution by service

Table 3 Descriptive statistics: Frequency of episodes per policyholder among periods

Set	Service	Mean	Median	St. Dev	Skewness	Kurtosis	Min	Max	Q(0.75)	Q(0.95)
A	DC	0.07	0.00	0.32	6.04	49.75	0.00	7.00	0.00	1.00
	OH	0.05	0.00	0.21	4.43	20.88	0.00	2.00	0.00	0.00
	RPC	0.02	0.00	0.24	17.43	469.87	0.00	15.00	0.00	0.00
	SC	0.02	0.00	0.23	17.37	416.53	0.00	12.00	0.00	0.00
	SMV	0.23	0.00	0.78	4.69	32.32	0.00	16.00	0.00	2.00
B	DC	0.05	0.00	0.26	5.99	49.01	0.00	5.00	0.00	0.00
	OH	0.04	0.00	0.19	4.91	25.76	0.00	3.00	0.00	0.00
	RPC	0.01	0.00	0.19	18.93	491.25	0.00	8.00	0.00	0.00
	SC	0.02	0.00	0.19	17.07	379.47	0.00	7.00	0.00	0.00
	SMV	0.17	0.00	0.64	5.18	38.45	0.00	14.00	0.00	1.00
C	DC	0.06	0.00	0.27	5.66	44.21	0.00	5.00	0.00	0.00
	OH	0.05	0.00	0.21	4.41	20.68	0.00	2.00	0.00	0.00
	RPC	0.02	0.00	0.20	16.02	318.78	0.00	7.00	0.00	0.00
	SC	0.02	0.00	0.20	15.19	296.58	0.00	6.00	0.00	0.00
	SMV	0.19	0.00	0.67	4.64	30.11	0.00	10.00	0.00	2.00

to each other in sets A and C, where as the rest of the groups in the A set exhibits longer tails.

By comparing severity charts in Fig. 3, where zero episodes data are excluded, we can state that are no evident changes in the shape before, during, and after COVID-19. This is expected as the pandemic has a more significant impact on frequency than expenditure. Among the most significant changes in severity, we observe the increase in the A set's leptokurtic shape (e.g., DC, OH, and RPC), though no longer tail between
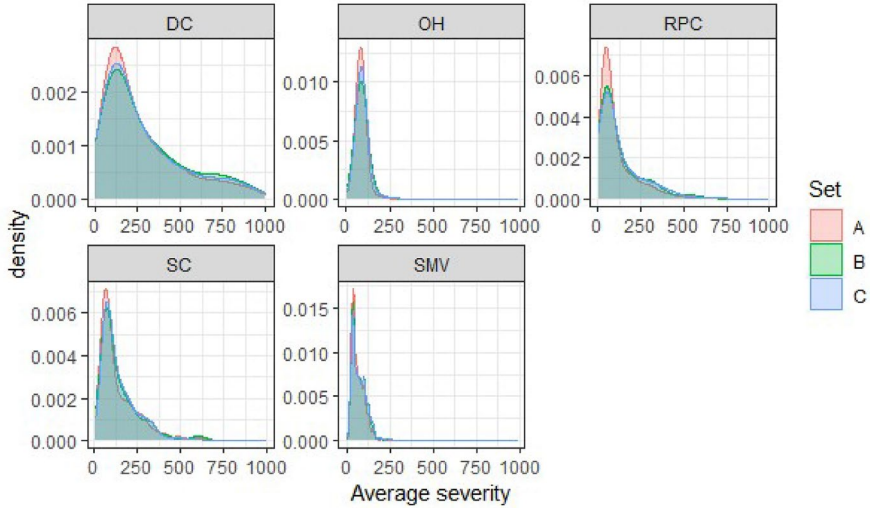
Fig. 3 Average Cost distribution by group pf services

the examined periods. Descriptive statistics in Table 4 confirms the mean increase and the kurtosis decrease from set A to C.

The increase in the mean can be justified by a large number but less expensive medical-care episodes in the pre-pandemic periods. Where as, as stated before, post-pandemic episodes, not strictly necessary and probably less costly, can be postponed.

Table 4 Descriptive statistics: Episode severity per policyholder among periods

Set	Service	Mean	Median	St. Dev	Skewness	Kurtosis	Min	Max	Q(0.75)	Q(0.95)
A	DC	472.25	230.00	788.84	8.98	182.42	6.00	24,000.00	510.00	1,700.00
	OH	85.97	85.00	36.48	8.12	145.10	7.50	1,100.00	95.00	120.00
	RPC	120.51	62.50	157.38	5.14	53.72	5.95	2,698.00	142.50	399.20
	SC	199.35	96.00	534.31	11.91	199.18	8.00	13,000.00	200.00	470.58
	SMV	67.52	51.96	114.09	24.69	756.89	2.90	4,400.00	86.44	130.00
B	DC	502.33	251.00	747.89	5.42	54.13	10.00	12,330.00	600.00	1,676.30
	OH	92.53	90.00	47.23	7.90	103.71	8.00	958.00	100.00	145.00
	RPC	134.09	70.00	155.44	3.13	19.22	3.00	1,500.00	180.00	400.00
	SC	214.09	102.46	505.50	7.65	68.01	7.75	5,576.92	204.17	549.80
	SMV	77.62	61.38	146.44	21.49	560.55	2.30	4,900.00	100.00	148.00
C	DC	493.22	250.00	750.91	6.06	70.92	5.00	13,990.00	600.00	1,700.00
	OH	92.39	90.00	40.19	7.08	111.61	6.50	978.00	100.00	145.00
	RPC	138.31	70.00	156.93	2.60	12.03	8.50	1,089.00	180.00	425.00
	SC	225.77	102.46	590.54	7.25	57.52	12.90	5,499.99	200.00	411.17
	SMV	81.61	63.08	193.53	22.82	635.11	2.30	7,500.00	100.00	150.00

Lastly, to deeply explore the frequency where policyholder behavior seems to have changed, we analyze the quarterly probability mass at zero, corresponding to the event of no episodes.

The box plot in Fig. 4 summarizes the distribution of the quarterly no-episode frequency observed during the (twelve) quarters included in set A. It displays its median and first and third quartiles in the gray box. At the same time, the dots and triangles points refer to the four quarters in sets B and C, respectively. As expected, each group's highest no-episode frequency is registered during Q2-2020 (i.e., during lockdown). However, the 2021 data shows an alignment to the pre-pandemic data as the triangle data points are most included between the first and third quartiles of the set A. Therefore, we can conclude that in the post-pandemic period analyzed, COVID-19 appears to have had no significant effect on the behavior of policyholders with the private medical coverage provided by a HIP.

3.2 Tweedie results

The second part of our analysis is devoted to assess the capability of a standard actuarial approach, based on a Tweedie GLM, to predict the HIP expenditure during and after the pandemic. To this aim, we conduct the analysis by splitting our dataset into a training set, composed of the A set, and a test set composed by the union of the set B and C. The training dataset represents 65,13% of the number of services and 62.92% of the total expenditure, where as the test set the 34.87% and 37.08%, respectively. The analysis is carried out implementing five independent Tweedie models, one for each group of services. As reported in Eqs. (2.6) and (2.7), we consider gender, age, quarter, and year as covariates; age and year are quantitative and continuous, where as gender and quarters are categorical variables. It is worth noting, that the parameters are estimated considering only the observation of the training set, in order to predict values on the test set.
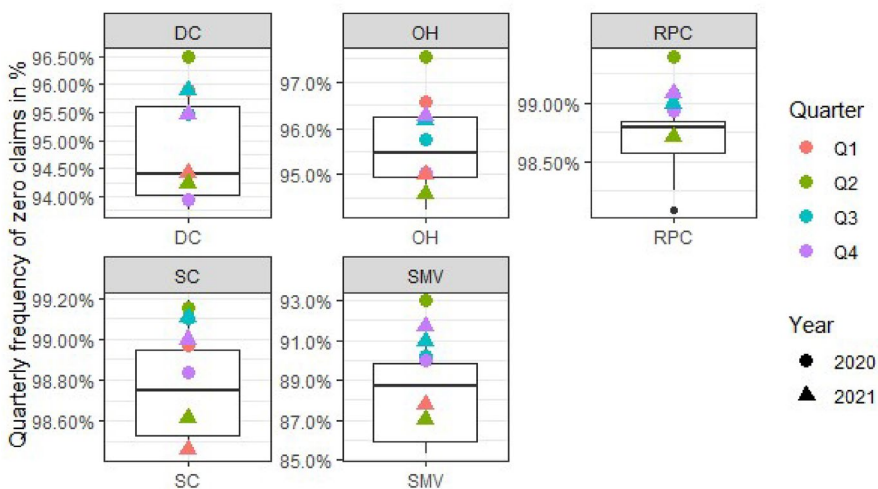

Fig. 4 Box plot of quarterly no-episode frequency during A period. Filled circles and triangles represent the quarterly no-episode frequency during a pandemic and post-pandemic quarters in different colors
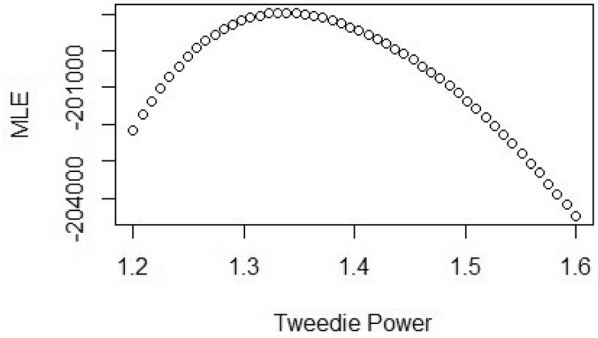
Fig. 5 Maximum Likelihood Estimate

Table 5 Output of tweedie regression

Parameter	Group of Services				
	DC	OH	RPC	SC	SMV
Intercept	−628.73 (*)	−1303.6 (***)	572.34 ( )	−381.67 ( )	−69,851.87 ( )
log(Year)	83.26 (*)	171.75 (***)	−75.37 ( )	50.05 ( )	–
Year	–	–	–	–	51.89 ( )
Year 3	–	–	–	–	0 ( )
Q2	0.03 ( )	0.08 (**)	0.11 ( )	0.17 ( )	-0.02 ( )
Q3	−0.28 (***)	−0.3 (***)	−0.38 (***)	−0.6 (***)	−0.46 (***)
Q4	0.12 (*)	0.01 ( )	−0.14 ( )	0.00 ( )	−0.25 (***)
Male	−0.23 (***)	−0.14 (***)	−0.47 (***)	−0.38 (***)	−0.63 (***)
Age	−0.28 (*)	−0.24 (**)	−0.08 ( )	0.08 ( )	−0.16 (*)
Age 2	0.01 (**)	0.01 (***)	0.01 ( )	0.00 ( )	0.00 (**)
Age 3	0.00 (***)	0.00 (***)	0.00 ( )	0.00 ( )	0.00 (***)

Signif. codes: 0 (***) 0.001 (**) 0.01 (*) 0.05 (.) 0.1 ( ) 1 − Not Used

The first stage of the calibration of the Tweedie GLM consists of the selection of the optimal value of the power p, which, for our purposes, needs to vary between 1 to 2. The methodology we adopt consists of testing a sequence of p between 1 and 2 by running iterative models and selecting the value corresponding to the maximum log-likelihood value. In Fig. 5, we show the traditional log-likelihood values that show an inverse "U" shape.

In Table 5, we show the parameter estimates and the significance measured by p-values.

The probability distribution of r.v. $_kS_h$ will be estimated conditioned to each risk class h, the latter being defined by combining the levels of the covariates considered here. All the insured belonging to the same risk class have the same estimates in terms of expected and tail behavior.

Figure 6 shows the distribution of the expected expenditure $E(_kS_h)$ (solid line) per age, group of services, and year, predicted via Tweedie GLM, compared to the observed values (solid points). Moreover, we plot the 99% confidence interval estimate of our predictions, by adding shadowed areas to our lines.

The graphs illustrate that the expenditure shape in each group remains essentially the same during and after the pandemic. It means that there is no evidence of a change in the policyholder behavior per age. Anyway, by looking at the left column of the graphs
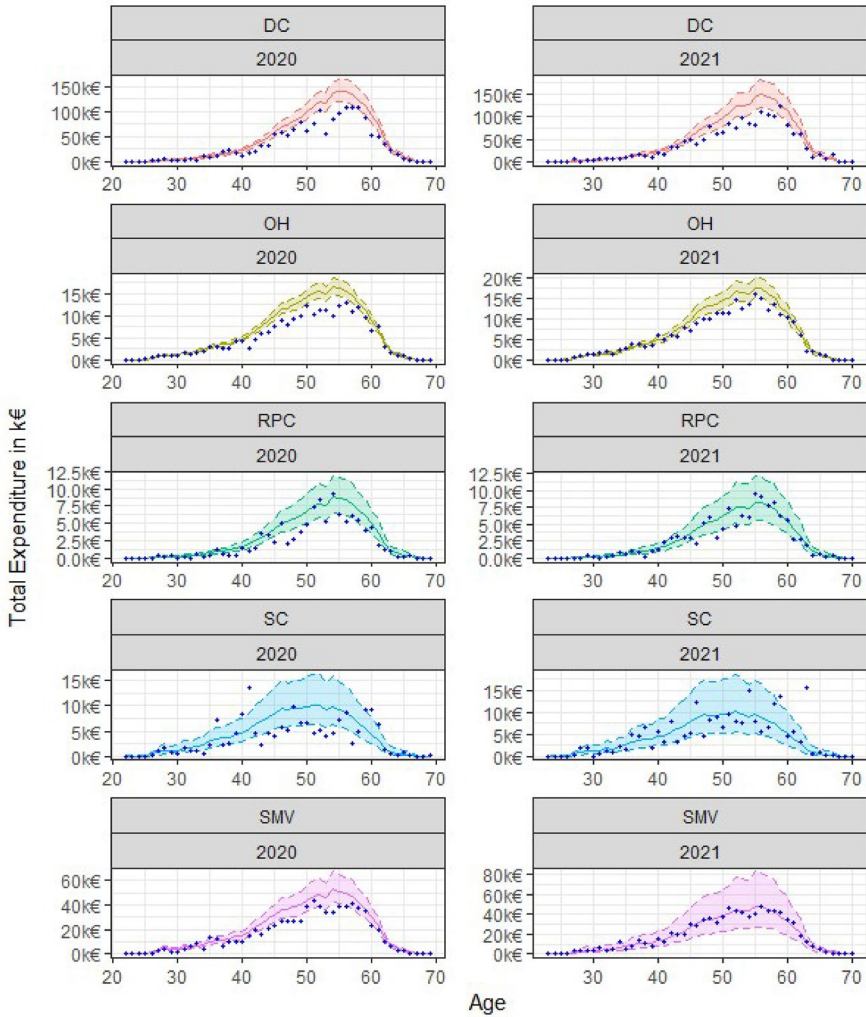
Fig. 6 Health care expenditure: a comparison of observed (solid points) and predicted value (solid line) with confidence interval (shadowed area) per Age, Group of services and Year

(i.e., the year 2020), there is a reduction in the expenditure in all groups, given the decrease in the frequency of the episodes, as previously exposed in Fig. 4 and Table 3. Moreover, the observed data (solid points) are generally located on or under the lower confidence interval of the mean value with a confidence level of 99%. These values confirm the exceptional nature of the pandemic phenomenon observed in 2020. Otherwise, the right column refers to the predicted value in 2021 by age, where the observed values are mainly in the prediction area. By way of example, SMV is the group with the highest number of episodes and the second in total expenditure (see Table 2) and displays observed values strictly in line with the predicted one. Similar behavior can be viewed in SC and RPC groups, where as DC and OH seem to overpredict the observations slightly.

Considering that the Tweedie model proposed also provides the predicted values concerning gender, quarter, and year, we compare in Fig. 7 the time series of the observed values with the predicted one.

The outcomes provide a clear picture of the expenditure trend during and after the pandemic and the model's prediction capability. Panels (a) and (b) figure the trend over different groups of services and quarters for females and males, respectively; the graphs limit the past data to 2019 for simplicity. As observable, there is no firm evidence of a difference between females and males in the observed data, which means that the medical services provided in each group are not limited to gender, as could be in the case of gynecology.
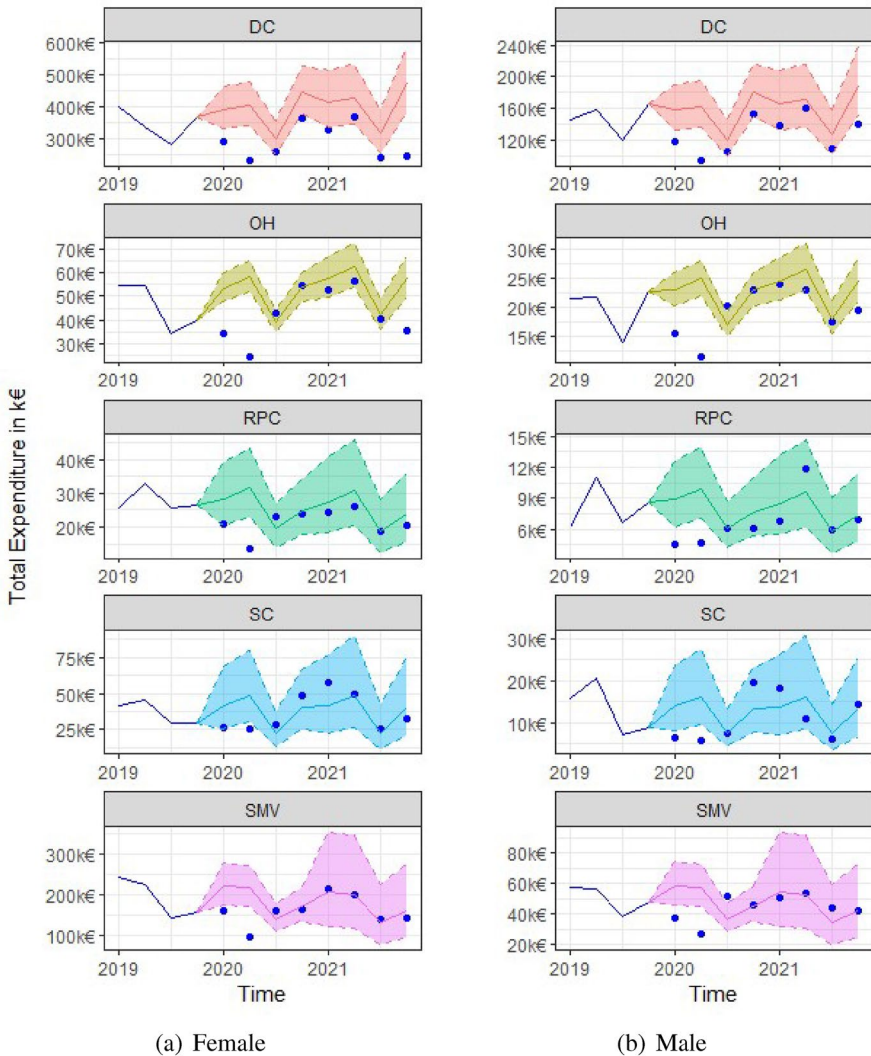

(a) Female   (b) Male

Fig. 7 Health care expenditure: time series of observed (solid points) and predicted value (solid line) with confidence interval (shadowed area) per Age, Gender, Group of services and Quarter

On the contrary, quarterly observed data, especially in the first two quarters of 2020, clearly depicts the model's limits to predict the effect of lockdown measures adopted by the Italian government. However, the rest of the aforementioned quarterly data allows us to capture almost all the observed data. This demonstrates, on the one hand, the excellent prediction capacity of the model. On the other hand, after the government's restrictive measures, the situation, at least for HIP, has returned to normal. The total expenditure predicted by the Tweedie model for years included in the test data is summarized in the following Table 6.

The yearly expenditure foreseen by the Tweedie model in each group of services shows that the 2021 values are generally included in the confidence interval except for DC, where the cost exceeds the lower bound of about 5%. Anyway, the goal of the HIP is to estimate the overall expenditure accurately. In this context, the model allows a difference between the predicted and observed mean values of about 15% and, in any case, well within the confidence interval. This result confirms the ability of the Tweedie model to predict healthcare cost data.

4 Conclusion

This study investigates the effects of COVID-19 on the health care expenditure of an Italian private HIP, which provides five groups of medical-care services. We first analyze if a relevant change in policyholder behavior is detected after the pandemic. To this aim, we adopt a basic actuarial approach based on a frequency-severity method, and we compare the before-pandemic data with the during and after one. Our results confirm no relevant evidence of a change in policyholder behavior. The only exception is observed during the pandemic due to the lockdown measure adopted by the Italian government in the first two quarters of 2020. In 2021 the main effect was a slight decline in the number of episodes but substantial stability of the average expenditure.

Then, we apply a Tweedie regression model whose main advantage is to analyze the distribution of the loss cost or pure premium, representing the total expenditure per policyholder. This distribution is particularly indicated considering the large number of zero episodes, we observe in the data for each group of services. This particular case of GLM is used to test the

Table 6 Total expenditure per Year, observed vs predicted values in ,000

Year	Service	Observed	Upper	Mean	Lower
2020	DC	1,616.69	2,573.77	2,164.10	1,821.15
	OH	226.86	329.27	292.70	260.30
	RPC	102.35	189.58	136.58	98.68
	SC	167.79	339.69	205.39	124.93
	SMV	742.26	1,199.39	946.26	746.64
	Total	*2.855,96*	*4.631,70*	*3.745,03*	*3.051,71*
2021	DC	1,728.46	2,843.63	2,284.76	1,837.35
	OH	270.18	363.90	313.85	270.80
	RPC	121.20	199.21	131.69	87.32
	SC	213.85	381.97	202.92	108.45
	SMV	881.06	1,515.35	880.14	511.22
	Total	*3.214,75*	*5.304,05*	*3.813,36*	*2.815,13*

capacity to predict health care expenditure even in stressful conditions. We conduct a quantitative analysis comparing observed values with point estimates and confidence interval estimates obtained by the regression model. Furthermore, this analysis is carried out concerning basic policyholder features such as gender and age to asses if some profiles have been most affected by the pandemic. We highlight that, despite the presence of lockdown restrictions, the predictive model captures the observed values well, confirming that the post-pandemic situation returns to a state of normality of behavior. This conclusion is approved by the fact that the values observed in 2021 are generally within the confidence interval produced by the model on pre-pandemic data. Furthermore, the differentiated analysis by gender and age confirms no clear evidence of a change in policyholder behavior in each group of services analyzed.

We conclude that using Tweedie GLM to examine the HIP expenditure is a valid forecasting method. Its dynamics over time confirm an excellent ability to detect potential behavior changes.

Declarations

Conflicts of interest The authors declare they have no financial interests. The authors equally contributed to the study conception and design.

References

Banthin, J., Simpson, M., Buettgens, M., Blumberg, L.J., Wang, R.: Changes in Health Insurance Coverage Due to the COVID-19 Recession: Preliminary Estimates Using Microsimulation. Quick Strike Series, Urban Institute (2020)

Bundorf, M.K., Gupta, S., Kim, C.: Trends in US Health Insurance Coverage During the COVID-19 Pandemic. JAMA Health Forum. **2**(9), e212487 (2021)

Daykin, C.D., Pentikainen, T., Pesonen, M.: Practical Risk Theory for Actuaries. Chapman and Hall/CRC, Boca Raton, FL (1993)

Frees, E.: Regression Modeling with Actuarial and Financial Applications (International Series on Actuarial Science). Cambridge University Press, Cambridge (2009). https://doi.org/10.1017/CBO9780511814372

Frees, E., Jin, X., Lin, X.: Actuarial Applications of Multivariate Two-Part Regression Models. Ann. Actuar. Sci. **7**(2), 258–287 (2013). https://doi.org/10.1017/S1748499512000346

Frees, E.: Frequency and Severity Models. In: Frees, E., Derrig, R., Meyers, G. (eds.) Predictive Modeling Applications in Actuarial Science (International Series on Actuarial Science, pp. 138–164. Cambridge University Press, Cambridge (2014)

Frees, E.W., Lee, G., Yang, L.: Multivariate frequency-severity regression models in insurance. Risks **2016**(4), 4 (2016). https://doi.org/10.3390/risks4010004

Hansen, J., Reinecke, A., Schmerer, H.J.: (2021). Health Expenditures and the Effectiveness of Covid-19 Prevention in International Comparison, CESifo Working Paper Series 9069, CESifo

Kaas, R.: (2005). Compound Poisson distribution and GLM's - Tweedie's distribution in Proceedings of the Contact Forum "3rd Actuarial and Financial Mathematics Day", pages 3-12. Brussels: Royal Flemish Academy of Belgium for Science and the Arts

Kurz, C.F.: Tweedie distributions for fitting semicontinuous health care utilization cost data. BMC Med. Res. Methodol. **17**, 171 (2017). https://doi.org/10.1186/s12874-017-0445-y

Tweedie, M.C.K.: Some statistical properties of Inverse Gaussian distributions. Virginia J. Sci. New Ser. **7**, 160–165 (1956)

Zuo, F., Zhai, S.: The influence of China's COVID-19 treatment policy on the sustainability of its social health insurance system. Risk Manag. Healthc. Policy **2021**(14), 4243–4252 (2021)