Modelling the impact of Covid-19 pandemics on health insurance associated services demand

A. Fernández-Fontelo[1], P. Puig[1], M. Guillén[2], D. Moriña[2]

[1] Departament de Matemàtiques, Universitat Autònoma de Barcelona (UAB), Cerdanyola del Vallès, Spain, amanda@mat.uab.cat
[2] Department of Econometrics, Statistics and Applied Economics, Universitat de Barcelona (UB), Riskcenter-IREA, Barcelona, Spain

Abstract

Health insurance is one of the branches of insurance with the greatest penetration in the Spanish market; the same happens in many developed countries. The claim rate in health insurance has suffered the impact of the Covid-19 pandemic in 2020 and 2021, especially regarding consultations and medical events that could be postponed. Mobility restrictions led to a decline in the use of insurance by policyholders and a transformation of the interaction between patients and health workers with greater use of telephone consultation. This work aims to study how to determine if, (i) due to the effect of postponing visits or (ii) due to the consequences of having suffered the virus (persistent Covid or side effects), there will be an excess of claims and, where appropriate, when a claiming increase will occur.

1.   Introduction

In the last months, there has been a big global concern around the 2019-novel coronavirus (SARS-CoV-2) infection, prompting the World Health Organization (WHO) to declare a public health emergency in early 2020. The pandemic induced by this virus has significantly impacted many aspects of human activity. In addition to the direct consequences concerning deaths caused by the Covid-19 infection and the saturation of health systems in many countries (including Spain and neighboring countries), a decrease in the use of health services has been detected in both the public health system and services associated with private health insurances in 2020. With more than 12 million insured, health insurance is one of the insurance branches with the highest penetration in the Spanish market. More than 25% of the population has this type of coverage, exceeding 35% in some areas (UNESPA, 2020). In 2020 and 2021, the Covid-19 pandemic impacted the companies' claim rates, especially regarding medical consultations and actions that were apparently of low priority and thus were postponed. The mobility restrictions led to a decline in the use of insurance services, as well as a transformation of the interaction between patients and health workers, with a larger use of telephone consultations. The concern is whether there will be an excess of claims in 2022 and later years, either due to postponing visits or due to the pandemic consequences (e.g., because of persistent Covid symptoms or secondary effects). In the public health system, there is already evidence of an increased frequency of use of health services. However, it is not straightforward to determine if the higher frequency of claims that will be observed will be equal to or greater than the infra-loss rate that was observed during the pandemic period. For this purpose, we present here a new method for

misreporting (i.e., under-reporting or over-reporting) estimation, which builds on Fernández-Fontelo et al. (2016, 2019). We use this method to determine: (i) whether the rebound effect occurs uniformly for all health insurance coverages or only for certain health insurance coverages, (ii) whether the rebound effect occurs homogeneously or varies depending on the insured characteristics, and (iii) to estimate the time point at which the initial benefit level is recovered. Some analysis will be conducted on the effects of health spending at the public system level, but the implications for private health insurance will also be of interest. Above all, it is expected that in order to monitor the effects of the pandemic in the coming years, these forms of analysis –such as the one presented in the current work– will be used as population groups with different sociodemographic characteristics or impacts on the use of health services cannot be directly compared.

This work aims to estimate and evaluate the pandemic's impact on health insurance, estimating the degree of under-reporting of health insurance usage, mainly in 2020, as well as the degree of over-reporting of health insurance usage nowadays as a result of the pandemic consequences. In addition, we want to develop a system that monitors claims to detect changes in the dynamics of medical insurance use in particular and any other branch in general. Although the main focus of the current research is on the development of a new method for misreporting estimation, which has also been empirically tested using simulated data, this method can also be applied to aggregated (anonymized) data from the health portfolio. We believe that the results and conclusions derived in the current research can be extended to other branches, as well as used to assess potential inequalities between countries or regions. In 2020, it was estimated that the total benefits provided by health insurance in Spain reached 6,300 million euros, of which 6,200 million correspond to the provision of medical services. In 2019, it was estimated that the total benefits provided by this type of insurance was 6,600 million euros, of which 6,500 million correspond to the provision of medical services.

The current work is organized as follows. Section 2 presents the model, its properties, and a method for estimating the model's parameters. Section 3 shows the main results of a preliminary simulation study to evaluate how the model behaves in practice in both the under-reporting and over-reporting scenarios.

2.   The model

Let X_n be the following stationary INAR(1) process: $X_n = \alpha \circ X_{n-1} + Z_n$, where $\alpha \circ X_n$ is the so-called binomial thinning operator and $Z_n \sim \mathrm{Poisson}(\lambda)$. Then, X_n follows a Poisson distribution with $\mathbb{E}(X_n) = \mu_X = \lambda/(1-\alpha) = \sigma_X^2 = \mathbb{V}(X_n)$. Suppose that the processes X_{n-1} and Z_n are independent for all n. In addition, the auto-covariance and auto-correlation functions of this process are respectively $\gamma(k) = \mathrm{Cov}(X_n, X_{n+k}) = \alpha^k \sigma_X^2$ and $\rho(k) = \mathrm{Cor}(X_n, X_{n+k}) = \alpha^k$.

Let Y_n be an observed and potentially misreported (i.e., under-reported or over-reported) process such that:

$$Y_n = \begin{cases} X_n & 1 - \omega \\ \vartheta \diamond X_n & \omega, \end{cases} \tag{1}$$

2

where $\vartheta \diamond X_n$ is the so-called fattering-thinning operator defined as:

$$[\vartheta \diamond X_n | X_n = x_n] = \sum_{j=1}^{x_n} W_j, \tag{2}$$

and W_j are independent and identically distributed (i.i.d) random variables distributed following the probability mass function (pmf) defined below:

$$\mathbb{P}(W_j = k | \phi_1, \phi_2) = \begin{cases} 1 - \phi_1 - \phi_2 & \text{if } k = 0 \\ \phi_1 & \text{if } k = 1 \\ \phi_2 & \text{if } k = 2 \\ 0 & \text{otherwise,} \end{cases} \tag{3}$$

with $\vartheta = (\phi_1, \phi_2)$. Therefore, the observed process Y_n coincides with the true process X_n with probability $1 - \omega$; otherwise, the observed process Y_n is either lower (under-reported) or higher (over-reported) than the true process X_n with probability ω. The parameter $0 < \omega < 1$ is the frequency of the misreporting phenomenon, and parameters ϕ_1 and ϕ_2 indicate whether our process is under-reported or over-reported (i.e., if $\phi_1 + 2\phi_2 > 1$ the process is over-reported; if $\phi_1 + 2\phi_2 < 1$ the process is under-reported). Note that if $\phi_1 = 1$, the process is not misreported. In addition, a less flexible version of the operator defined in expression (2) can be derived if $W_j \sim \text{Binomial}(2, \phi)$. Although the distribution in expression (3) is the most easy way to model over-reporting in count time series, other probability distributions with no compact support (e.g., the Poisson or Geometric distributions) can be taken into account here as well.

The operator defined in expression (2) follows a 2nd-order Hermite distribution with parameters $\mu_X \phi_1$ and $\mu_X \phi_2$, where $\mu_X = \lambda/(1 - \alpha)$ is the expectation of X_n and $\phi_j = \mathbb{P}(W = j)$ for $j = 1, 2$. We can see the above using the idea of probability generating functions (pgf) as follows: First, the marginal distribution of the observed process $X_n \sim \text{Poisson}(\mu_X)$ with pgf $G_X(s) = e^{\mu_X(s-1)}$. Second, the pgf of the random variable W in expression (3) is given by $G_W(s) = (1 - \phi_1 - \phi_2) + \phi_1 s + \phi_2 s^2$. Finally, the pgf of the operation defined in expression (2) is given by $G_{\vartheta \diamond X_n}(s) = G_X(G_W(s)) = \exp(\mu_X \phi_1(s - 1) + \mu_X \phi_2(s^2 - 1))$, which is the pgf of a 2nd-order Hermite distribution with parameters $\mu_X \phi_1$ and $\mu_X \phi_2$. Hence, the expectation and variance of the random variable $\vartheta \diamond X_n$ are $\mathbb{E}(\vartheta \diamond X_n) = \mu_X(\phi_1 + 2\phi_2)$ and $\mathbb{V}(\vartheta \diamond X_n) = \mu_X(\phi_1 + 4\phi_2)$, respectively.

We already now that the latent process is distributed as $X_n \sim \text{Poisson}(\lambda/(1 - \alpha))$ and the fattening-thinning operator follows $\vartheta \diamond X_n \sim \text{Hermite}(\phi_1 \mu_X, \phi_2 \mu_X)$, where $\mu_X = \lambda/(1 - \alpha)$. In addition, we can see that the distribution of the observed process Y_n is the following mixture of a Poisson distribution and a 2nd-order Hermite distribution:

$$Y_n = \begin{cases} \text{Poisson}(\mu_X) & 1 - \omega, \\ \text{Hermite}(\mu_X \phi_1, \mu_X \phi_2) & \omega, \end{cases} \tag{4}$$

with expectation and variance defined as $\mathbb{E}(Y_n) = \mu_X(1 - \omega(1 - \phi_1 - 2\phi_2))$ and

$$\mathbb{V}(Y_n) = \mu_X(1 - \omega(1 - \phi_1 - 4\phi_2)) + \mu_X^2 \omega(1 - \omega)(1 - \phi_1 - 2\phi_2)^2.$$

Finally, the auto-correlation function of the observed process Y_n is defined as:

$$\gamma(k) = \text{Cov}\,(Y_n, Y_{n+k}) = \mu_X \alpha^k \left(1 - \omega\left(1 - \phi_1 - 2\phi_2\right)\right)^2, \tag{5}$$

which can be derived computing $\mathbb{E}(Y_n, Y_{n+k})$, $\mathbb{E}(Y_n)$ and $\mathbb{E}(Y_{n+k})$. In fact, we already know $\mathbb{E}(Y_n) = \mu_X(1 - \omega(1 - \phi_1 - 2\phi_2))$ (and similarly for $\mathbb{E}(Y_{n+k})$ given that Y_n is a stationary process). In addition, $\mathbb{E}(Y_n, Y_{n+k})$ can be computed as a function of $\mathbb{E}(X_n, X_{n+k}) = \text{Cov}(X_n, X_{n+k}) + \mathbb{E}(X_n)\mathbb{E}(X_{n+k}) = \alpha^k \sigma_X^2 + \mu_X^2$. The auto-correlation function can be thus computed as:

$$\rho(k) = \text{Cor}\,(Y_n, Y_{n+k}) = \frac{\text{Cov}(Y_n, Y_{n+k})}{\sqrt{\mathbb{V}(Y_n)}\sqrt{\mathbb{V}(Y_{n+k})}} = \tag{6}$$
$$= \frac{\alpha^k \left(1 - \omega\left(1 - \phi_1 - 2\phi_2\right)\right)^2}{(1 - \omega(1 - \phi_1 - 4\phi_2)) + \mu_X \omega(1 - \omega)(1 - \phi_1 - 2\phi_2)^2} =$$
$$= \alpha^k c(\alpha, \lambda, \omega, \phi_1, \phi_2).$$

Note that $c(\alpha, \lambda, \omega, \phi_1, \phi_2)$ is constant with respect to k, and hence the ACF of the observed process Y_n is proportional (up to a multiplicative constant) to that of the latent process X_n (i.e., the ACF is geometrically decreasing at rate k given that $0 < \alpha < 1$). Note also that we take always by definition that $\rho(0) = 1$.

Model inferences

Parameters of the model (i.e., $\alpha, \lambda, \omega, \phi_1$ and ϕ_2) are estimated using the likelihood function, but this is not directly tractable as described by Fernández-Fontelo *et al.* (2016, 2019). According to these authors, we can compute the likelihood function here using the forward algorithm, which is based on the so-called transition and emission probabilities. In particular, the transition probabilities are given by the conditional pmf of the stationary INAR(1) process defined as:

$$\mathbb{P}(X_n = x_n | X_{n-1} = x_{n-1}) =$$
$$= \sum_{j=0}^{\min(x_n, x_{n-1})} \binom{x_{n-1}}{j} \alpha^j (1 - \alpha)^{x_{n-1}-j} \mathbb{P}(Z_n = x_n - j), \tag{7}$$

where $\mathbb{P}(Z_n = x_n - j)$ is computed in the following with the pmf of a Poisson(λ) distribution. In addition, the emission probabilities here are defined as given:

$$\mathbb{P}(Y_n = y_n | X_n = x_n) = \begin{cases} 0 & \text{if} \quad x_n < y_n/2 \\ (1 - \omega) + \omega p_n & \text{if} \quad x_n = y_n \\ \omega p_n & \text{if} \quad x_n > y_n \\ \omega p_n & \text{if} \quad x_n < y_n,\ x_n \geq y_n/2, \end{cases} \tag{8}$$

and probabilities p_n are computed using the following recursive relation:

$$p_n = \frac{1}{n(1 - \phi_1 - \phi_2)} \left[\phi_1 (x_n - (n-1)) p_{n-1} + \phi_2 (2x_n - (n-2)) p_{n-2} \right], \quad (9)$$

which was computed following Baena-Mirabete and Puig (2020). Note that we consider under-reporting when $x_n > y_n$ and over-reporting when $x_n < y_n$, but $x_n \geq y_n/2$. Therefore, if there is over-reporting, our model assumes that what we see (i.e., y_n) is at most twice what actually happens (i.e., x_n).

Finally, with the transition and emission probabilities in expressions (7) and (8), the likelihood function of the observed process Y_n is recursively computed using the forward algorithm, which is essentially based on the forward probabilities. That is, the likelihood function is thus calculated as:

$$\mathbb{P}\left(Y_{1:N} = y_{1:n}\right) = \sum_{x_N = \frac{y_N}{2}}^{\infty} \gamma_N \left(y_{1:N}, x_N\right), \quad (10)$$

where the forward probabilities $\gamma_n \left(y_{1:N}, x_N\right)$ are defined as:

$$\gamma_n(y_{1:n}, x_n) = \mathbb{P}(Y_n = y_n | X_n = x_n)$$
$$\sum_{x_{n-1} = \frac{y_{n-1}}{2}}^{\infty} \mathbb{P}(X_n = x_n | X_{n-1} = x_{n-1}) \gamma_{n-1}(y_{1:n-1}, x_{n-1}), \quad (11)$$

considering $\mathbb{P}(X_1 = x_1) \sim \mathrm{Poisson}(\lambda/(1-\alpha))$. Note that expression (11) is based on the transition probabilities and emission probabilities, that in our model here are defined as given in expressions (7) and (8) respectively. In practice, we use numerical optimization procedures like the `nlm` function in R.

3. Preliminary simulated results

In order to evaluate our model capabilities for under-reporting and over-reporting detection and estimation, we have first simulated two time series, one with over-reporting and another with under-reporting. Table below provides the maximum likelihood point estimates for each of the parameters for both time series, as well as the standard errors. Our model appropriately identifies whether the time series is over-reported or under-reported, and the true values of the parameters are always contained in the 90% Wald confidence intervals. Note that, although we here know which time series is over- and under-reported, our model also provides an easy mechanism for identifying which misreporting phenomenon is present in our data. In particular, if $\phi_1 + 2\phi_2 > 1$, then we very likely have over-reporting. Contrarily, if $\phi_1 + 2\phi_2 < 1$, we may have the opposite case of under-reporting. Finally, if $\phi_1 + 2\phi_2 = 1$, we do not have over-reporting or under-reporting.

We also compared several theoretical and empirical moments for both simulated time series, such as the mean, variance, and the first auto-correlation coefficients. More precisely, for the over-reported time series, we observed a mean and variance of 7.00 and 13.65, respectively, while the corresponding theoretical values are 6.91

and 14.31. With respect to the first auto-correlation coefficients, we observed $\hat{\rho}(1) = 0.192$, $\hat{\rho}(2) = 0.126$, and $\hat{\rho}(3) = 0.037$, while the corresponding theoretical values are $\rho(1) = 0.2183$, $\rho(2) = 0.0655$ and $\rho(3) = 0.0196$. For the under-reported time series, the empirical mean and variance were 3.34 and 7.52 compared to the theoretical ones that were 3.40 and 6.82. The first three coefficients of the empirical auto-correlation were here $\hat{\rho}(1) = 0.163$, $\hat{\rho}(2) = 0.101$, and $\hat{\rho}(3) = 0.073$ compared to the theoretical ones that were $\rho(1) = 0.1433$, $\rho(2) = 0.0717$ and $\rho(3) = 0.0358$.

	Over-reporting				
	α	λ	ω	ϕ_1	ϕ_2
true parameter	0.3	3.0	0.7	0.1	0.8
point estimate	0.3578	3.2684	0.5501	0.0771	0.8072
std. error	0.1054	0.7352	0.1155	0.0297	0.0829
	Under-reporting				
	α	λ	ω	ϕ_1	ϕ_2
true parameter	0.5	3.0	0.7	0.2	0.1
point estimate	0.5184	3.0586	0.7354	0.1952	0.0784
std. error	0.1554	0.8808	0.0890	0.0511	0.0339

References

UNESPA. Informe Estamos Seguros. Published online 2020:290.

Ministerio de Sanidad. Acuerdo del Consejo Interterritorial del Sistema Nacional de Salud de 24 de febrero de 2021 sobre la pandemia de la COVID- 19 y la prevención y el control del cáncer. Published online 2021.

Fernández-Fontelo A, Cabaña A, Puig P, Moriña D. Under-reported data analysis with INAR-hidden Markov chains. Stat Med. 2016; 35(26):4875-4890. doi:10.1002/sim.7026

Fernández-Fontelo A, Cabaña A, Joe H, Puig P, Moriña D. Untangling serially dependent underreported count data for gender-based violence. Stat Med. 2019;38(22):4404-4422. doi:10.1002/sim.8306

Baena-Mirabete S, Puig P. Computing probabilities of integer-valued random variables by recurrence relations. Statistics & Probability Letters 2020; 161: 108719. https://doi.org/10.1016/j.spl.2020.108719