Insurance Mathematics and Economics 126 (2026) 103175

Contents lists available at ScienceDirect

Insurance: Mathematics and Economics

journal homepage: www.elsevier.com/locate/ime

# Back to normal? a method to test and correct a shock impact on healthcare usage frequency data

David Moriña [a,b,c,*], Amanda Fernández-Fontelo [a], Montserrat Guillén [b]

[a] Department of Mathematics, Universitat Autònoma de Barcelona, Cerdanyola del Vallès, Spain
[b] Department of Econometrics, Statistics and Applied Economics, Riskcenter-IREA, Universitat de Barcelona, Barcelona, Spain
[c] Centre de Recerca Matemàtica (CRM), Cerdanyola del Vallès, Spain

ARTICLE INFO

*2000 MSC:*
62P05
62M10
62F15

*Keywords:*
Health services
Risk modelling
Poisson offset
Covid-19
Time series
Pre-processing

ABSTRACT

A method based on Bayesian structural time series is proposed to predict healthcare usage trends and to test for changes in the series levels during or after an abnormal year, such as that of the 2020 COVID-19 pandemic. Our method can also serve to calculate correction factors for frequency count data that can be integrated in a preprocessing step before undertaking a cross-sectional statistical analysis, and, in this way, the impact of a shock can be eliminated. Here, adjustments are derived for a large private health insurer in Spain from estimates of average healthcare usage. Median claims rate levels in 2020 were 15 % down on 2019 figures, but rose in 2021 and 2022, when the rate was 11 % and 8 % higher than in 2019, respectively. Once the shock correction is incorporated in the preprocessing step, our approach is shown to outperform traditional time series techniques. Healthcare insurance usage in Spain did not fully go back to normal levels (assuming that pre-pandemic values represent normality) in 2022, with the exception of some patient groups and specific medical services. Our method can be implemented in other areas of risk analysis when frequency counts are exposed to shocks and it allows estimating the difference in claims volume between real figures and those estimated, had the shock not occurred.

## 1. Introduction

Following the world-wide outbreak of COVID-19, one of the problems faced by health insurers was that historical records of the use of medical services - which usually serve for the yearly update of future premiums - have become heavily influenced by the pandemic consequences (Kim et al., 2022b; Xu et al., 2021; Cantor et al., 2022). After the COVID-19 pandemic, analysts might be tempted to delete the information for the years 2020 to 2022, and return to using 2019 as their baseline, on the grounds that once things are back to normal, 2019 should be everyone's reference year. This poses the question as to whether 2020 to 2022 data remain useful for future projections. If rescaling were feasible, then data following a shock year could still be analyzed, net, that is, of the shock effect; otherwise, data from shock and post-shock years may not be comparable to those from regular periods.

Here, we seek to measure the impact of the mandatory lockdown driven by the Spanish authorities as a COVID-19 pandemic mitigation action in terms of the frequency of use of medical insurance services. Moreover, we take advantage of these estimates to correct individual portfolio (the company's inforce of policies) data since 2020 for statistical modeling in the subsequent years, so that no data have to be discarded. In so doing, we draw on traditional time series decomposition methods by looking at weekly data, in order to balance temporal granularity and statistical stability. The lockdown effect can be identified and deducted from the original series in order to derive future trends net of the shock.

Our objectives are: 1) the quantification of the impact of a shock on weekly claim frequency rate series, 2) the determination as to whether a claim frequency series continues to show a significant impact of that shock, and 3) the rescaling of frequency data to eliminate the effects of the impact. These objectives are achieved by following the modelling framework described in the following sections.

We illustrate our method with data drawn from a Spanish private health insurance company. We seek to quantify the effect of the pandemic outbreak and lockdowns on the time series of claims, by identifying the size of that shock in 2020 and the return to the pre-pandemic level of 2019. To do so, we design a step-by-step procedure that can help health insurers decide whether or not their portfolio information continues to be affected by the pandemic shock. In this way, they will know if the claims data are suitable for use in their predictive modeling,

* Corresponding author.
*E-mail address:* david.morina@uab.cat (D. Moriña).

https://doi.org/10.1016/j.insmatheco.2025.103175
Received 1 August 2024; Received in revised form 3 October 2025; Accepted 4 November 2025
Available online 24 November 2025

or, should the data continue to present signs of a post-pandemic effect, they can determine the size of this distortion and how the data might be corrected to make them suitable for accurate premium calculations. The analytical procedure we propose can be conducted with all types of claim, as well as by subgroups defined by claimant characteristics, including sex and age, and by type of healthcare service.

## 2. Background

Many analysts realize that the years 2020, 2021, and 2022 can be deemed special in terms of health insurance provision. Moreover, retaining those stressful periods is a challenge for risk analysis (Biancalana and Baione, 2022). Besides the catastrophic effect those particular periods might have on insurability and contract design in all lines of business (Hartwig et al., 2020), the problem is identifying whether the fluctuations in claims frequency and severity seemed to stabilize. In this regard, a tool that quantitatively adjusts claim frequency to correct for imbalances in claims (both underuse and overuse) is particularly valuable. This allows future analyses to rely on recent data rather than reverting to 2019 data.

The consequences of having to reschedule medical treatment and preventive care in 2020, which may have led to an overuse of medical services in 2021, and possibly 2022, are still unclear (Wilensky, 2022). Some medical specialties have suffered the effects more than others, indicating that the post-pandemic impact may still be significant for some specific medical use records. During the pandemic, hospitals had to ride the various waves of the pandemic, but on-going cancer treatments, including chemotherapy and surgeries, were discontinued as little as possible. Yet, many non-urgent surgeries, preventative care appointments, and some diagnostic services were canceled, resulting in delayed treatment and an increased healthcare burden for patients with non-COVID related conditions.

In a case study of a large insurance company in Spain, we find that health insurance claims during the pandemic (2020) were about 15 % down on pre-pandemic figures (2019), while in 2021 and 2022, the usage of health care services increased but did not return to 2019 levels. There are, however, a number of exceptions to this general result. For instance, the series of frequency health care claims for people aged 60+ quickly returned to the 2019 average, whereas some services, such as visits to a general practitioner (GP) for all ages, were higher in 2022 than in 2019.

The pandemic and lockdowns caused an unprecedented disruption to the global health insurance market (Przybytniowski et al., 2022; Szczygielski et al., 2022). First and foremost, it obliged many people to stay home, so that the number of medical appointments, hospital admissions, and other medical services in 2020 were initially down on 2019 figures. The fall in the frequency of claims for healthcare usage can be attributed to the postponement of elective medical procedures, such as surgeries and other invasive treatments, which are typically covered by health insurance. These procedures were often put on hold due to the strain that the healthcare system was under, as well as the mobility restrictions in many areas. Likewise, the fear of contracting COVID-19 in healthcare settings led some to avoid seeking medical care, even for serious conditions that would typically require health insurance coverage. Overall, it has been reported that the decrease in health insurance claims during the pandemic was attributable to a combination of these factors (Plott et al., 2020), but that a rebound was expected after the pandemic. The resilience of health insurance systems to the impact of contagious diseases like COVID-19 has been discussed from a theoretical approach (Hong and Seog, 2023). From the methodological point of view, it is important to bear in mind that we can date quite precisely the moment when the pandemic consequences start impacting the usage of health services, so we can define the pre- and post-shock periods. Sometimes, the date a shock starts impacting the series of interest is not known and must be estimated. In these cases, the date could be estimated within the Bayesian framework following the ideas of Barry and Hartigan (1993) or Rosenberg and Young (1999).

The pandemic created a state of emergency across all areas of healthcare, resulting in reduced resources for the treatment of other illnesses. This pattern was similar in many countries around the globe (Xu et al., 2021; Mogharab et al., 2022), but there is evidence that some demographic groups were more badly affected than others. For example, older adults in the Netherlands (Mizee et al., 2022) and Germany (Michalowsky et al., 2021) suffered substantially higher cancellation rates of medical visits and the postponement of medical care in 2020 than younger adults. In this regard, several studies discuss the consequences of delaying the treatment of non-emergency diseases (Kim et al., 2022a; Kotrych et al., 2022; Di Martino et al., 2022) and scheduled surgeries (Ricciardiello et al., 2021).

In general, in many countries, private health insurance enjoyed something of a boost after 2020 as the market expanded, with a price hike in premiums and a considerable rise in the number of policy holders. One reason why premiums rose, despite the fact that the initial effect of COVID-19 was a reduction in claims, is that the pandemic changed how health insurance companies view pandemics as a real catastrophe rather than a potential risk (Richter and Wilson, 2020). Indeed, many insurers opted to take a more cautious approach to risk after COVID-19, resulting in higher premiums, increased deductibles, and reduced coverage. Additionally, some insurers have changed their eligibility criteria and now charge higher premiums for those with pre-existing conditions. The pandemic has also triggered an increase in the cost of healthcare services (Poisal et al., 2022), reflecting the higher demand, and increased costs for both medical supplies and the delivery of care. As a result, many health insurance companies have increased their prices. A further trigger has been the greater uncertainty surrounding a possible increase in claims due to the fact that some pathologies that might otherwise have been diagnosed, were not. In these instances the treatment is likely to be more expensive than had the disease been identified in its early stages. Furthermore, reports of the long-term effects of COVID-19 have also raised fears of a protracted increase in the level of claims.

## 3. Methods and data

Here, we propose an approach that adapts the well-known difference-in-differences (DD) methodology to the time series setting. We do so by explicitly modeling the counterfactual of a time series observed both before and after the occurrence of an event of interest, potentially impacting the evolution of the process, the case, in this instance, of the shock provoked by the COVID-19 pandemic. Our approach provides a fully Bayesian time-series estimate of the effect and uses model averaging to construct the most appropriate synthetic control for modeling the counterfactual that the shock never happened. The traditional DD approach is based on a static regression model that assumes independent and identically distributed data despite the fact that the design has a temporal component, and usually considers only two time points: before and after the shock. As discussed, for instance, in Antonakis et al. (2010), when fit to serially correlated data, static models yield overoptimistic inferences with too narrow uncertainty intervals. The challenges of DD schemes can be mitigated by employing state-space models combined with flexible regression techniques to capture the temporal dynamics of an observed outcome. State-space models differentiate between two components: the state equation, which governs the transition of latent variables over time, and the observation equation, which links the system's state to measurable data. This separation provides significant flexibility and power, as discussed in Leeflang et al. (2009) in the context of marketing research. The proposed method adopts three key aspects of the state-space framework. First, our method allows for flexible incorporation of various assumptions about latent states and emission processes, such as local trends and seasonality. Second, we apply a fully Bayesian approach to model the temporal progression of counterfactual scenarios and incremental impacts, offering versatility in how posterior inferences are summarized. Finally, the regression component avoids rigidly adhering to a fixed set of control

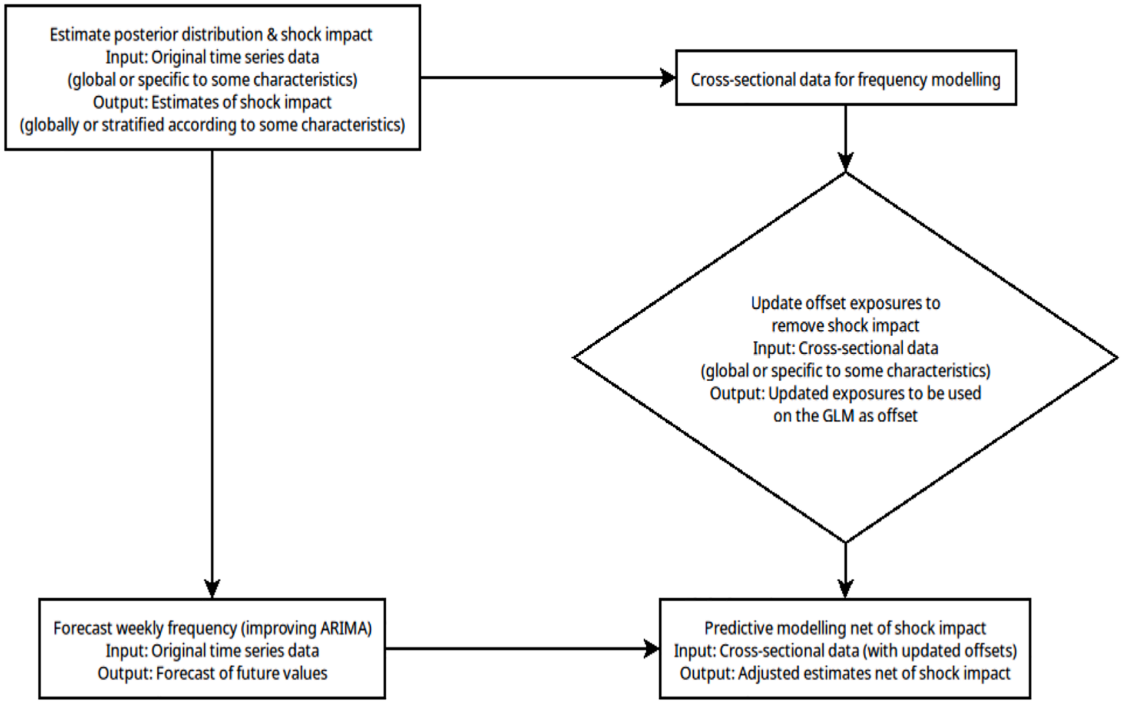


**Fig. 1.** Methodology for testing the impact of a shock and, afterwards, filtering the data.

variables by accounting for uncertainty in both the influence of individual predictors and the selection of relevant predictors, thus reducing the risk of overfitting. To summarize, this approach extends the commonly used DD approach to time-series analysis by explicitly modeling the counterfactual for a time series observed both before and after the intervention. It offers two key improvements over other existing methods: it provides a fully Bayesian estimate of the intervention's effect in a time-series context, and it employs model averaging to create the most suitable synthetic control for the counterfactual modeling (Brodersen et al., 2015).

By using the proposed approach, as summarized in Fig. 1, the expected claim frequency obtained in the cross-sectional analysis corresponds to what would be expected once the impact of the shock is removed.

Methodological frameworks based on BSTS have also been adopted in other research areas, notably for quantifying the effects of physical modifications on water consumption (Schmitt et al. (2018)) and for evaluating the role of sea rescue operations in shaping irregular migration dynamics (Rodríguez Sánchez et al. (2023)).

### 3.1. Shock impact detection and evaluation

Let $y_t$ denote the observation $t$ in a real-valued time series $t = 1, \ldots, n$, where $n$ is the last period of observation. A structural time series (STS) model can be described by a pair of model equations relating $y_t$ to a $d$-dimensional vector of latent state variables $\alpha_t$:

$$y_t = Z_t'\alpha_t + \epsilon_t, \epsilon_t \sim N(0, H_t) \tag{1}$$

$$\alpha_{t+1} = T_t\alpha_t + R_t\eta_t, \eta_t \sim N(0, Q_t). \tag{2}$$

Eq. 1, or the observation equation, links observed data $y_t$ with the unobserved latent state $\alpha_t$; whereas Eq. 2, or the transition equation, defines how the latent state $\alpha_t$ evolves over time. A model that can be described by these two equations is known to be in state space form, which defines a large class of models including the autoregressive integrated moving average or ARIMA (Scott and Varian, 2014).

Note that $'$ denotes transposition, $y_t$ represents the evolution in the weekly claims rate associated with a private health insurance company, $Z_t$ is a $d$-dimensional output vector, $T_t$ is a $d \times d$ transition matrix, $R_t$ is a $d \times q$ control matrix, $\epsilon_t$ is a scalar observation error with noise variance $H_t$, and $\eta_t$ is a $q$-dimensional system error with a $q \times q$ state-diffusion matrix $Q_t$, where $q \leq d$. In our setting, in particular, the best fitting model considers a static intercept, a seasonal effect of period 52 (as we are dealing with weekly data) and the effects of holidays. In this case, the general model described in Eqs. 1 and 2 can be rewritten as

$$y_t = c + \alpha_t + \epsilon_t, \epsilon_t \sim N(0, \sigma^2_{\epsilon,t}) \tag{3}$$

$$\alpha_{t+d} = -\sum_{i=0}^{s-2} \alpha_{t-i\times d} + \eta_t, \eta_t \sim N(0, \sigma^2_{\eta,t}), \tag{4}$$

where $c$ is a constant value, $s$ is the number of seasons and $d$ is the seasonal duration (number of time periods in each season, set to 1 for the week cycles but specifically defined for the holidays effects). The model can be thought of as a regression on $s$ dummy variables representing $s$ seasons and $\alpha_t$ denotes their joint contribution to the observed response $y_t$. The mean of $\alpha_{t+d}$ is such that the total seasonal effect is zero when summed over $s$ seasons.

Let $\theta$ denote the set of all model parameters and $\alpha = (\alpha_1, \ldots, \alpha_m)$ denote the full state sequence. Following Brodersen et al. (2015), a prior distribution $p(\theta)$ is specified on the model parameters jointly with a distribution $p(\alpha_0 \mid \theta)$ on the initial state values. We then sample from $p(\alpha, \theta \mid y)$ using a Markov chain Monte Carlo (MCMC) algorithm. To estimate the impact of a shock on the evolution of a time series based on the methodology proposed, draws of the model parameters $\theta$ and the state vector $\alpha$ given the observed data $y_1, \ldots, y_n$ in the training period are simulated. We denote $y_1, \ldots, y_n$ as $y_{1:n}$. The posterior simulations are then used to simulate from the posterior predictive distribution $p(\tilde{y}_{n+1:m} \mid y_{1:n})$ over the counterfactual time series $\tilde{y}_{n+1:m}$ given the observed pre-shock activity $y_{1:n}$. Finally, we use the posterior predictive samples to compute the posterior distribution of the pointwise impact $y_t - \tilde{y}_t$ for each $t = (n+1), \ldots, m$ and the posterior distribution of cumulative impact, summarized by their median, which can be interpreted as the shock level, and its 95 % credible intervals (CI). When using the Bayesian methodology described to estimate the parameters, the method is usually referred to as a Bayesian structural time series (BSTS). A different approach, based on classical time series models and intervention analysis was used in Li et al. (2017) to explore the increase in the number of terrorist bombing attacks observed by late 2011, after United States' troops withdrawal from Afghanistan and Iraq. The Bayesian approach requires specifying a prior distribution on the model parameters and on the initial state values. We used the default prior distributions defined in the R package *CausalImpact* (Brodersen et al., 2015). For the model defined in Eqs. 3 and 4, the priors for the inverse of the variances $\sigma^2_{\epsilon,t}$ and $\sigma^2_{\eta,t}$ are $\mathcal{G}(10^{-2}, 10^{-2}s^2_y)$, where $s^2_y = \frac{1}{n-1}\sum_t (y_t - \bar{y})^2$ and $\mathcal{G}(a,b)$ is the Gamma distribution with expectation $\frac{a}{b}$. The prior for the initial state $p(\alpha_0 \mid \theta)$ is a flat normal distribution. More details on the modelling approach and prior choices can be found in Brodersen et al. (2015). TODO: Additionally,

As shown below, this methodology is also capable of providing accurate forecasts. Its performance is compared to that of usual time series modelling, such as ARIMA (see for instance Shumway and Stoffer (2017)) by means of two commonly used error measures - that is, the root mean square error (RMSE) and the mean absolute percentage error (MAPE) -, computed according to Eqs. 5 and 6 respectively. We assume that $n^*$ periods are forecast, and we denote the observed values $O_j$ and the forecast values $F_j$, with $j = 1, \dots, n^*$.

$$RMSE = \sqrt{\frac{1}{n^*} \cdot \sum_{j=1}^{n^*} (O_j - F_j)^2}, \tag{5}$$

$$MAPE = \frac{100}{n^*} \sum_{j=1}^{n^*} \left| \frac{O_j - F_j}{O_j} \right|. \tag{6}$$

In order to select the best fitting ARIMA model, Akaike's Information Criterion (AIC) as defined below will be used.

$$AIC = 2 \cdot k - 2 \cdot \ln(\hat{L}), \tag{7}$$

where $k$ is the number of estimated parameters in the model and $\hat{L}$ is the maximized value of the likelihood function for the model.

### 3.2. Implications for pre-processing data net of a shock impact in GLM frequency modelling

The BSTS-estimated median shock level (the values reported in Table 1) can be used afterwords to pre-process data in order to filter out the effect of the impact of the pandemic shock in a cross-sectional analysis including individual data on observed claims and claimant covariate information. The details on how to use the BSTS-estimates in a cross-sectional analysis are provided below.

Let us introduce the notation of cross-sectional data at time $T$. Suppose that $N_{iT}$ is the observed number of claims for policy holder $i$ in year $T$, where $i = 1, \dots, N_T$ and $N_T$ is the number of policy holders in year $T$. Assume that a vector of $K+1$ characteristics $x_{iT}$ is available for individual $i$ in year $T$ where the first component is a constant and $\nu_{iT} \in (0,1]$ denotes exposure, i.e. the fraction of the whole year when the policy was valid, then predictive modeling of the expected number of claims usually establishes that $E(N_{iT}|x_{iT}, \nu_{iT}) = \nu_{iT}\mu(x_{iT})$, where $\mu$ is a function that links the effect of the covariates to the expected number of claims, provided that this effect is proportional to exposure (Wüthrich and Merz, 2023). In the Poisson model, $\mu(x_{iT}) = \beta' x_{iT}$, with $\beta = (\beta_0, \dots, \beta_K)$ a vector of parameters to be estimated, where $\beta_0$ is an intercept.

If the impact of a shock has to be eliminated from these data for the purpose of predictive modeling and making comparisons between years, then the expected number of claims should be affected by a change and this can easily be achieved through the exposure component. Therefore new individual exposures $\nu^*_{iT}$ can be defined so that

$$\nu^*_{iT} = \nu_{iT} \times m_T$$

where $m_T = 1 + f_T$ is obtained from the Bayesian estimate described in the previous section, which can be expressed as a percent change, $f_T$. For instance, a global correction could be applied to the Poisson model by considering $f_T = -0.15$ as given in Table 1 (a percent reduction of 15 %), and therefore $m_T = 1 - 0.15 = 0.85$, so individuals with a valid policy in the whole year ($\nu_{iT} = 1$) using the considered health services in 2020 are only 85 % exposed to use these services compared to individuals observed in 2019 ($\nu^*_{iT} = 1 \times 0.85$).

Note that $f_T$ refers to the cumulated posterior distribution median relative change. A positive $f_T$ means that the shock has produced an excess of counts, corresponding to increased exposure. A negative $f_T$ means that the shock has reduced the number of counts, corresponding to decreased exposure. Note that $f_T$ may be a global correction that is equal for all individuals, or it may have been calculated by subgroups, so that it could change by individual. We have opted to drop the $i$ subindex in order to eliminate complex notation from the methodology.

Once the corrected exposures, $\nu^*_{iT}$, are used in the Poisson model estimation process, the model outcome is net of the pandemic shock impact. In simpler terms, the estimation proceeds by first fitting a model to the 2019 data in order to approximate how many individuals would have used health insurance-related services in 2020 had the COVID-19 pandemic not occurred. This model is then used to produce counterfactual forecasts for 2020, which are compared with the actual observations. The resulting discrepancy is interpreted as the shock impact. Finally, the observed 2020 data are adjusted by removing this estimated impact, thus providing shock-free values that can be employed in the cross-sectional models. Below we illustrate how this can be implemented.

In this illustration, we have incorporated the Bayesian estimates into the individual exposures and fitted Poisson regression models for claims frequency in a simple example with portfolio data sets before, during, and after the pandemic shock. We compare analyses of single vs multiple-year predictive modeling, with and without the inclusion of the shock correction, and show that predictive claim models for pre- and post-pandemic yearly portfolios are actually comparable. Some additional results, including the comparison of the distribution of the predicted number of claims per individual with and without the pandemic effect, reveal a persistent change in the distribution. The analysis by groups of individuals, also shows that processing before eliminating the pandemic effect need not be homogeneous for the whole portfolio. It is important to note that the impact of shocks cannot be directly estimated in the Poisson model, since the data used for this model are typically aggregated by year, sex, and age group.

### 3.3. Gap of cost estimation

If average claim cost information is available, the approach proposed here can be used to estimate what would annual costs have been in the absence of the shock. Therefore, an estimate of the gap between the real observed cost and the estimated cost level without the shock can be provided.

In our case, if we assign an average unit cost of $C$ per visit or health care claim, costs can be estimated. We need to predict number of policies with 0 claims under the difference scenarios of shock removal. Similarly, the number of policies with $1, 2, 3, 4, 5+$ claims can be estimated, respectively. These approximations can be carried out based on the Poisson models that were estimated with the data that were already preprocesses to eliminate the pandemic shock effect. Other more sophisticated models can also be used for predictive modeling purpose, because the only requirement here is to use the suitable preprocessed data or the original data.

Finally, the gap of cost is the difference between the real observed cost and the estimated cost level obtained from the claims predictions without the shock.

### 3.4. Data

We observed the weekly number of medical claims associated with a private health insurance scheme run by one of the largest companies in

Spain from the beginning of 2019 to the end of 2022, by type of service, province, sex and age group. The overall monthly number of contracts was also taken into account, to control for observed trends unrelated to the consequences of pandemic claims. To make the latter weekly, we performed a linear interpolation process and the final considered outcome was the density of claims by 1,000 active contracts, computed as $y_t = \frac{c_t}{n_t} \cdot 1000$, where $c_t$ is the number of claims produced at week $t$ and $n_t$ is the total number of active contracts at week $t$.

A weekly frequency count corresponds to the number of claims reported to the insurance company in the course of one week. A claim refers to the use of a medical service, such as a visit to a GP, a blood test, an x-ray, surgery, a visit to a specialist, the use of an ambulance or treatment in an emergency room or hospital that was covered by the insurance policy.

The insurer providing us with their data was affected by the enforcement of social distance measures, adopted in Spain as in many other countries. These triggered a sudden fall in the frequency of health insurance claims and this decrease, moreover, was substantial given that during the lockdowns in 2020 people reduced their normal activities and canceled preventive medical visits, while medical services were primarily focused on treating COVID-19 patients. In addition, the insurance company's portfolio initially fell markedly in 2020, affected in all probability by the impact of the death of a certain number of policy holders during the pandemic. Then, as in other European countries, the portfolio grew substantially owing to the saturation of the public health system. The latter situation raised grave concerns among many citizens who opted to underwrite private health insurance to protect themselves and their families in case of a global medical service collapse. Although private hospitals, accessible to holders of a health insurance contract, were also relatively saturated during the pandemic, COVID-19 mainly impacted public hospitals.

The immediate impact of the COVID-19 pandemic can be quantified by comparing the number of claims in 2019 and 2020, while the medium- and long-term consequences can be quantified by comparing the number of claims in 2019, with those in 2021 and 2022. To illustrate the ability of the methodology proposed to quantify the impact of the pandemic on the use of private health services, in what follows we describe in detail the evolution in medical claims globally, by sex, among those aged over 60 and in the three Spanish provinces that are home to the country's largest cities (that is, Madrid, Barcelona, and Valencia). A number of specific services of particular interest are also analyzed.

The R code used to generate the results and figures described below is available in the GitHub repository https://github.com/dmorinya/BSTS_HealthInsurance. The results reported in Tables 2 and 3 can be reproduced with the cardiology subset, available in https://data.mendeley.com/datasets/kb9fyth3xh/1. The rest of our data cannot be made publicly available due to owner restrictions. All coding has been done in R software (Team, 2019), using the packages *bsts* (Scott, 2024) and *CausalImpact* (Brodersen et al., 2015).

## 4. Results

Although the impact of the COVID-19 pandemic onset on the usage of health insurance services could be calculated as the reduction in claims frequency observed in 2020 with respect to 2019 and the increase observed in 2021 and 2022, also with respect to 2019, this information could not then simply be used to forecast the future behavior of claims. Indeed, in the case study we conduct here, we could simply state that a total of 11,754,419 claims (recall this corresponds to the whole portfolio of a large insurance company, where a claim is defined as any medical service provided, including for example a blood test) were made in 2019, while 10,305,088 were made in 2020, 13,337,053 in 2021 and 14,129,286 in 2022. That is, a decrease of 12 % was recorded in 2020 and increases of some 13 and 20 % were recorded, respectively, in 2021 and 2022, compared in both cases to 2019. However, on the basis of these figures, data analysts cannot then predict future usage trends. Moreover, these figures are affected by the number of policy holders in the portfolio, in particular, an increase of new policy holders from 2020 onwards.

A more sophisticated approach, and one that would allow analysts to make a forecast of future behavior, is provided by the classical ARMA time series models, although the original process is not stationary (Kwiatkowski - Phillips - Schmidt - Shin (KPSS) test p-value lower than 0.01). In order to deal with a stationary series, one regular difference is applied (KPSS test p-value of the differenced series is 0.1). When fitting the best ARMA model to the differenced series, based on AIC, to the period January 2019 to June 2022 - that is an ARMA(1, 1) process with a seasonality of 52 weeks - the weekly frequency of claims forecast produced for the period July 2022 to December 2022 is shown in Fig. 2.

It is evident that the ARMA-based forecast is only capable of capturing the central tendency of the process, but at the price of considerable discrepancies between observed and forecast values (RMSE of 52,731.67 and MAPE of 18.55 %). The forecast for the same period provided by the BSTS approach outlined above is shown in Fig. 3. Several BSTS alternative models were considered (using a static intercept, seasonal, autoregressive and local linear trend state component) and the one that provided lowest RMSE and MAPE was preferred, corresponding to a static intercept and seasonal state component. All the reported results in this section correspond to this approach. Here, it is clear that the forecast produced is much more accurate than that provided by the classical approach. Indeed, in this case, the RMSE is around 33,870 and the MAPE is around 18.47 %; thus, in both instances the error is lower than when adopting the classical ARMA-based approach. The BSTS approach is also capable of providing a visual idea of how the series would have behaved in the counterfactual situation of no COVID-19 shock, represented by the solid blue line in Fig. 3.

Table 1 reports the estimated effect - both the median and 95 % credible intervals (instead of the mean and mean-related dispersion measures, as is common in the Bayesian context, although reported results do not differ substantially from the corresponding posterior means) - of the pandemic shock on the overall weekly time series of claims frequency rate observations and by group of policy holders, geographical area and medical specialty, by comparing the indicated annualities. Independent models were fitted stratifying by group of policy holders, geographical area and medical specialty, without any covariates. As these estimates are capturing the impact of the shock on the annualities comparison, could be used as an individual exposure adjustment in a cross-sectional analysis of the claims frequency as detailed on Section 3.2.

As can be seen, a quite remarkable reduction was recorded in the usage of health insurance services in 2020 compared to the previous year (with the exception of general medicine services), while a boost is evident in most cases in 2021, most notably in general medicine, cardiology and osteopathy. The non significant 1.2 % decrease estimated for general medicine in 2020 (and a part of the 14 and 21 % increases reported, respectively, for 2021 and 2022) can be attributed to the fact that some of the COVID-19 testing was conducted by this service. Geographically, the behavior of Barcelona and Valencia is largely similar in both comparisons, while Madrid records a greater reduction in usage in 2020 and smaller increases in both 2021 and 2022. This might reflect the fact that Madrid has an older exposure profile; yet, this hypothesis cannot be confirmed with the data available as the monthly number of contracts is aggregated only at the national level. The behavior of both sexes is also very similar, but the reduction in usage among individuals aged over 60 in 2020 is significantly greater than that of the general population. Moreover, in this group, no increase in usage was recorded in 2021 (indeed, we find a reduction of around 8 %, although it is non-significant, and a decrease of around 3 % in 2022, again non-significant). Here, it should be borne in mind that COVID-19-associated mortality rates were much higher among this subpopulation; thus, it is reasonable to expect that the older adults (> 60) that survived are healthier than in the past and that their need for health services is not as great as that of older adults in 2019.

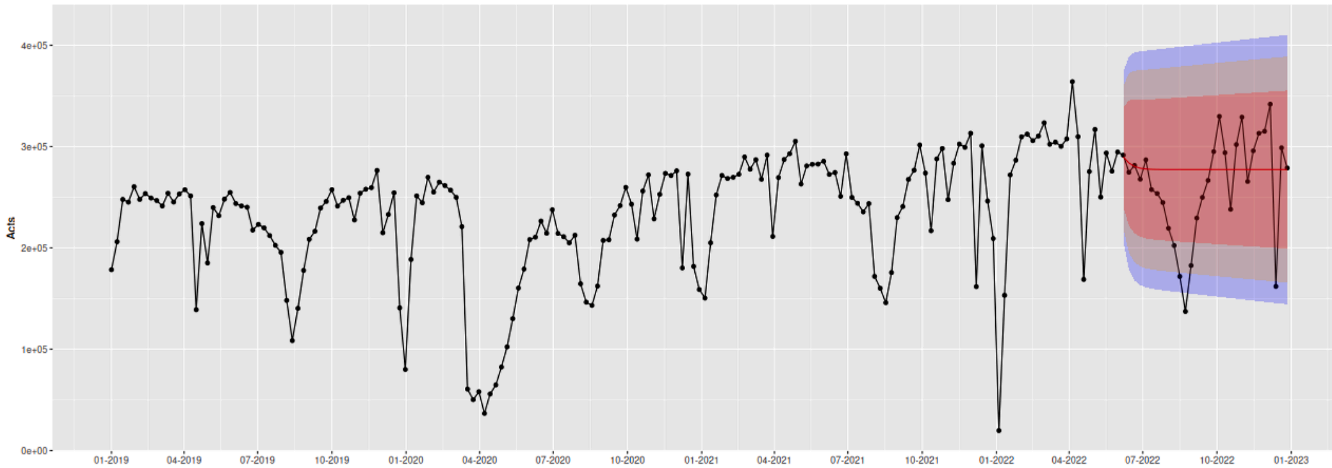


**Fig. 2.** Number of weekly claims (black) and the ARIMA-based forecast for the period July-December 2022 (red), 95 % confidence bands in light blue, 90 % confidence bands in orange and 75 % confidence bands in light red. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.).

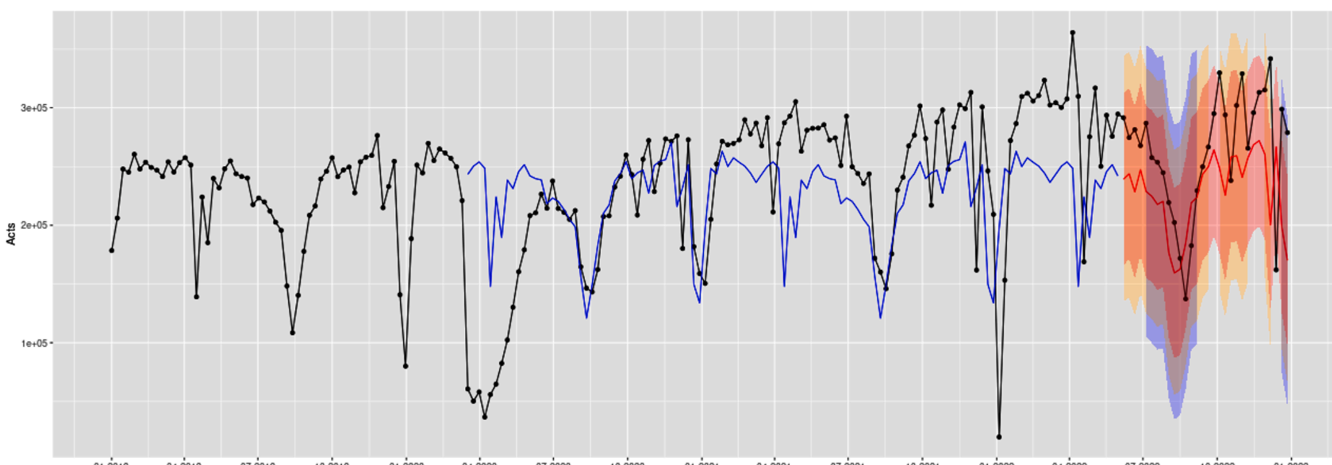


**Fig. 3.** Number of weekly claims (black) and BSTS forecast for the period July-December 2022 (red), 95 % confidence bands in light blue, 90 % confidence bands in orange and 75 % confidence bands in light red. The blue line depicts the prediction under the counterfactual situation of no COVID-19 shock. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.).

The results are robust to different choices of prior distributions. In particular, when wide normal priors are applied, the estimates remain virtually unchanged, as shown in Table S9 (Supplementary Material), and are consistent with those reported in Table 1.

To illustrate how the Bayesian estimates provided by the proposed methodology can be incorporated in data analysis in a standard cross-sectional context, we next fit several Poisson regression models to the number of cardiology claims (i.e. a visit to a cardiologist) by age (30 – 60 / > 60) and sex (male/female). The distribution of these variables is reasonably balanced (49.99 % females, 36.09 % over 60). We fit one model per year (2019, 2020, 2021, and 2022) and a global model including data for all the years under consideration. Additionally, we fit a Poisson regression model using an offset correction accounting for the global Bayesian estimates reported in Table 1 (i.e. -15 % for 2020, 11 % for 2021, and 8.1 % for 2022) and a further Poisson regression model using specific Bayesian estimates for each subgroup of age and sex. To summarize, the results of the time series analysis provide the factor that is introduced as a data weighting to eliminate the excess counts. This is a common procedure when conducting data analyses from data arising from complex sampling schemes that require using sample weights (see Pfeffermann (1993) or Kish and Frankel (1974) for instance). The estimates yielded by each of these models are shown in Table 2. The differences are immediately evident, especially in the case of the impact of the age group parameter when using specific corrections. Note that this model reflects the expected claims net of the impact of the pandemic. Thus, when all years are combined and the shock is filtered out, the incidence in adults aged > 60 is 1.91 times greater than that in adults aged 30 to 60, while the incidence in men is 17 % higher than that in women, see the last row of Table 2. A year-by-year analysis without any shock correction reveals lower incidence rates in general, but in particular for 2022.

Table 3 shows that the observed data are best fitted by the Poisson regression model without any correction, whereas those incorporating global and specific corrections model (with less and more detail, respectively) the counterfactual that the impact of the COVID-19 pandemic is negligible. This indicates that insurance companies can use all the information to compare the temporal behavior of usage series, without having to exclude any specific year.

Moreover, note that the last row of Table 3 shows what the outcome distribution would have been in the absence of the pandemic shock. Indeed, the full sample model with this specific correction reports the expected behavior net of the 2020 pandemic shock, and comparison with the uncorrected full sample reveals how the estimated incidences impact the final premiums.

**Table 1**
BSTS estimated median percentage change attributable to impact of lockdown and post-pandemic behavior over the frequency usage of health insurance associated services.

| | 2019–2020 | | 2019–2021 | | 2019–2022 | |
|---|---|---|---|---|---|---|
| | Difference (95 % CI) | p-value | Difference (95 % CI) | p-value | Difference (95 % CI) | p-value |
| Total | -15.0 % (-19.0 %, -11.0 %) | < 0.01 | 11.0 % (2.8 %, 20.0 %) | < 0.01 | 8.1 % (0.9 %, 16.0 %) | 0.01 |
| Females | -15.0 % (-19.0 %, -10.0 %) | < 0.01 | 11.0 % (2.4 %, 20.0 %) | < 0.01 | 7.8 % (0.5 %, 16.0 %) | 0.02 |
| Males | -15.0 % (-19.0 %, -11.0 %) | < 0.01 | 11.0 % (3.3 %, 21.0 %) | < 0.01 | 8.5 % (1.4 %, 16.0 %) | < 0.01 |
| Over 60 | -22.0 % (-26.0 %, -18.0 %) | < 0.01 | -8.2 % (-16.0 %, 0.3 %) | 0.03 | -3.0 % (-10.0 %, 4.8 %) | 0.21 |
| Oncology | -14.0 % (-19.0 %, -8.3 %) | < 0.01 | 3.7 % (-3.4 %, 12.0 %) | 0.17 | 3.0 % (-3.0 %, 9.7 %) | 0.17 |
| Cardiology | -13.0 % (-17.0 %, -8.4 %) | < 0.01 | 14.0 % (5.0 %, 24.0 %) | < 0.01 | 12.0 % (4.1 %, 21.0 %) | < 0.01 |
| Obstetrics | -15.0 % (-19.0 %, -10.0 %) | < 0.01 | 8.6 % (0.4 %, 18.0 %) | 0.02 | 6.4 % (-1.0 %, 15.0 %) | 0.05 |
| Urology | -14.0 % (-18.0 %, -9.0 %) | < 0.01 | 15.0 % (5.9 %, 25.0 %) | < 0.01 | 13.0 % (4.8 %, 22.0 %) | < 0.01 |
| General medicine | -1.2 % (-5.9 %, 4.1 %) | 0.30 | 14.0 % (7.0 %, 23.0 %) | < 0.01 | 21.0 % (13.0 %, 30.0 %) | < 0.01 |
| Osteopathy | -26.0 % (-30.0 %, -21.0 %) | < 0.01 | 15.0 % (3.3 %, 28.0 %) | < 0.01 | 24.0 % (14.0 %, 36.0 %) | < 0.01 |
| Madrid | -19.0 % (-23.0 %, -14.0 %) | < 0.01 | 2.5 % (-5.8 %, 12.0 %) | 0.30 | 3.2 % (-4.2 %, 11.0 %) | 0.22 |
| Barcelona | -12.0 % (-17.0 %, -7.1 %) | < 0.01 | 18.0 % (8.3 %, 28.0 %) | < 0.01 | 12.0 % (3.6 %, 21.0 %) | < 0.01 |
| Valencia | -12.0 % (-16.0 %, -7.0 %) | < 0.01 | 21.0 % (13.0 %, 31.0 %) | < 0.01 | 17.0 % (8.9 %, 26.0 %) | < 0.01 |

**Table 2**
Parameter estimates of incidence rate ratios for Poisson regression modeling of the frequency of claims originating in a cardiology service (p-values correspond to testing incidence rate equal to 1) for policy holders aged 30 + . A correction refers to a change in exposure before estimating the Poisson regression using the BSTS impact estimate for 2020, 2021, and 2022 either globally, the same for all individuals, or by gender/age group. A correction eliminates the impact of the pandemic shock.

| Model | $e^{\hat{\beta}_0}$ (p-value) | $e^{\hat{\beta}_{\text{Sex (Male)}}}$ (p-value) | $e^{\hat{\beta}_{\text{Age (60+)}}}$ (p-value) | Sample size |
|---|---|---|---|---|
| Only 2019 | 0.60 (< 0.01) | 1.15 (< 0.01) | 1.74 (< 0.01) | 75,218 |
| Only 2020 | 0.57 (< 0.01) | 1.14 (< 0.01) | 1.68 (< 0.01) | 73,477 |
| Only 2021 | 0.64 (< 0.01) | 1.15 (< 0.01) | 1.63 (< 0.01) | 86,737 |
| Only 2022 | 0.66 (< 0.01) | 1.15 (< 0.01) | 1.59 (< 0.01) | 91,188 |
| Full sample (without correction) | 0.62 (< 0.01) | 1.15 (< 0.01) | 1.65 (< 0.01) | 326,620 |
| Full sample (with global correction) | 0.61 (< 0.01) | 1.15 (< 0.01) | 1.66 (< 0.01) | 326,620 |
| Full sample (with specific correction) | 0.56 (< 0.01) | 1.17 (< 0.01) | 1.91 (< 0.01) | 326,620 |

**Table 3**
Observed and expected distribution of number of cardiology claims in each Poisson regression model considered. Period 2019–2022.

| Model | 0 claims | 1 claim | 2 claims | 3 claims | 4 claims | 5 or more claims |
|---|---|---|---|---|---|---|
| Observed | 150,670.0 | 120,040.0 | 38,283.0 | 11,499.0 | 3,861.0 | 2,267.0 |
| Full sample (without correction) | 150,562.9 | 113,790.0 | 45,585.7 | 13,008.4 | 2,977.5 | 695.5 |
| Full sample (with global correction) | 152,588.6 | 113,316.4 | 44,666.6 | 12,559.7 | 2,836.2 | 652.5 |
| Full sample (with specific correction) | 156,958.7 | 110,657.2 | 42,913.5 | 12,371.3 | 2,966.2 | 753.1 |

If cost information is available, the approach proposed can be used to estimate annual costs in the absence of the shock. In our case, if we assign an average unit cost of 50€ per visit to the cardiology service, Table 3 shows that the accumulated cost of cardiology visits over the period 2019–2022 is around 12,975,500€. Actual annual costs were 3,020,850€ in 2019; 2,704,550€ in 2020; 3,489,900€ in 2021, and 3,760,200€ in 2022. Using the global adjustment, the total cost in the absence of the pandemic can be estimated at 12,825,174€ which implies an excess cost equal to 150,326€ (i.e. 12,975,500€–12,825,174€). When incorporating the specific age/gender correction, the estimated excess cost is equal to 413,169€. Note that these figures are already accommodated to the size of the portfolio.

The usage made of other services, including that of general medicine, was also analyzed as described above for cardiology, and the results were very similar to those reported in this section.

## 5. Discussion

The method we propose is able to estimate the impact of the COVID-19 pandemic/lockdowns on the usage of services covered by private health insurance. Our analysis of a Spanish insurance portfolio shows a marked fall-off in usage in 2020 that is largely independent of the medical service or geographical area. For example, the reduction is most evident in less urgent services, such as osteopathy, but in the case of more critical services - most notably cardiology and oncology - this reduction is clearly more limited by the nature of the associated diseases (Table 1). More interestingly, our model is also able to estimate the subsequent shock attributable to the pandemic. Here, we report a general increase in service usage in general (up to 21 % and 17 % in Valencia in 2021 and 2022 respectively), less clear when looking at specific services or geographic areas, probably because the increase is more subtle than the decrease during the lockdown and the time series is too short. In this sense, it would be interesting to analyze the behavior of the series when more recent data become available.

The described methodology provides insurance companies with a convenient alternative for processing their data before implementation of their pricing models. Indeed, by using our more realistic estimations, they no longer need ignore data for 2020 - the usual approach employed to date - which also means overlooking the subsequent overuse of health services in the wake of the COVID-19 pandemic. Similarly, the methodology also highlights for the healthcare services the differential behavior of the oldest population, reflecting the well documented “high costs of dying”, that is, the disproportionate expenditure on medical care at the end of life (Lubitz and Prihoda, 1984; Scitovsky, 2005).

One limitation is that we could not analyze the claims severity directly. However, since these data are mostly medical visits, a standard flat cost is usually applied. An additional consideration in this work is that one of the main objectives of the methodology in Brodersen et al. (2015) is to construct a time series that would have occurred without the intervention. In this line of thinking, plotting how the series would have looked without the pandemic intervention, as well as the pointwise differences and cumulative differences is interesting. We have omitted too many plots here because they are quite straightforward. Another potential limitation of the study is that the impact of the increase in the use of telemedicine visits during and after the pandemic has not been considered, as the study focuses on traditional medical services.

When dealing with time series data that represents count observations, it is crucial to exercise caution. Traditional Gaussian and linear time series models may not be suitable for such data unless the counts are sufficiently large to be approximated as continuous variables. For further insight, refer to Example 8.8.3 in Brockwell and Davis (2002). This consideration extends beyond the ARIMA model to include the BSTS model as well. For instance, Example 8 in the *bsts()* function documentation (Scott, 2024) highlights the inappropriateness of using the Gaussian family for count data. However, we have implemented our modelling effort to the weekly rate series, i.e. claims per policy.

## 6. Conclusion

When considering claim frequency data that might have been affected by a pandemic/lockdown shock, identification of the impact should help in understanding expected future trends. Here, we calculate an average yearly effect as a means of assessing the impact of the shock in 2021 and 2022 by medical service and subgroup of insureds. Our findings suggest that analysts need not be concerned by the impact of the pandemic on health insurance claim frequency after 2021 as it affects certain some specific groups. In our illustration, old adults and medical specialties would seem to have recovered pre-pandemic levels of usage (albeit that some services, most notably visits to general practitioners, present unusually high frequency values).

Our study highlights the need for regulators to encourage risk analyses that use good, preprocessed data when merging pre- and post-pandemic datasets so as to offset the misconception that such data are comparable, that shocks can be ignored and/or that they affect all individuals equally. Implementing a step-by-step analysis - like the one advocated here - should help in identifying the impact of a pandemic shock and even in clarifying if its effects have faded or whether they persist over time.

## Funding Statement

This research was funded by the Fundación MAPFRE (Becas Ignacio H. de Larramendi, 2021). A.F-F. thanks the Agencia Estatal de Investigación for financial support (IJC2020-045188I/AEI/10.13039/501100011033) and for a María Zambrano scholarship. M.G. and D.M. thank the Spanish Ministry of Science and Innovation, FEDER grant PID2019-105986GB-C21 and NextGenerationEU, grant number TED2021-130187B-I00. MG gratefully thanks the ICREA Academia Program. D.M. acknowledges support fom the Spanish Government through the Ministerio de Ciencia e Innovación (grant PID2022-137414OB-I00). This work is supported by the Spanish State Research Agency, through the Severo Ochoa and María de Maeztu Program for Centers and Units of Excellence in R&D (CEX2020-001084-M). We thank CERCA Programme/Generalitat de Catalunya for institutional support. We thanks support received by the Spanish Ministry of Science, Agencia Estatal de Investigación TED2021-130187B-I00/AEI/10.13039/501100011033 and ICREA Academia.

## Data Availability Statement

The R code used to generate the results and figures described below is available in the GitHub repository https://github.com/dmorinya/BSTS_HealthInsurance. The results reported in Tables 2 and 3 can be reproduced with the cardiology subset, available in https://data.mendeley.com/datasets/kb9fyth3xh/1. The rest of our data cannot be made publicly available due to owner restrictions.

## Ethical Standards

The research meets all ethical guidelines, including adherence to the legal requirements of the study country.

## CRediT authorship contribution statement

**David Moriña:** Conceptualization, Methodology, Data curation, Visualization, Writing - original draft; **Amanda Fernández-Fontelo:** Conceptualization, Methodology, Writing - original draft; **Montserrat Guillén:** Conceptualization, Writing - original draft.

## Declaration of competing interest

M.G. has received funds from insurance companies, but the funding organizations had no role in the design of the study; in the collection, analysis, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results. The authors declare no other potential conflicts of interest.

## Supplementary material

Supplementary material associated with this article can be found, in the online version, at 10.1016/j.insmatheco.2025.103175

## References

Antonakis, J., Bendahan, S., Jacquart, P., Lalive, R., et al., 2010. On making causal claims: a review and recommendations 21 (6), 1086–1120. https://doi.org/10.1016/j.leaqua.2010.10.010
Barry, D., Hartigan, J.A., 1993. A bayesian analysis for change point problems 88 (421), 309–319. https://doi.org/10.2307/2290726
Biancalana, D., Baione, F., 2022. A quantitative analysis on the effect of COVID-19 in a private health insurance plan expenditure. Quality & Quantity , 1–15, published online.
Brockwell, P.J., Davis, R.A., 2002. Introduction to time series and forecasting. Springer.
Brodersen, K.H., Gallusser, F., Koehler, J., Remy, N., Scott, S.L., 2015. Inferring causal impact using bayesian structural time-series models. Ann Appl Stat 9 (1), 247–274. https://doi.org/10.1214/14-AOAS788
Cantor, J., Sood, N., Bravata, D.M., Pera, M., Whaley, C., 2022. The impact of the COVID-19 pandemic and policy response on health care utilization: evidence from county-level medical claims and cellphone data. J Health Econ 82, 102581.
Di Martino, G., Cedrone, F., Di Giovanni, P., Romano, F., Staniscia, T., 2022. Impact of COVID-19 pandemic on oncological surgery activities: a retrospective study from a southern italian region. Healthcare 10 (11), 2329.
Hartwig, R., Niehaus, G., Qiu, J., 2020. Insurance for economic losses caused by pandemics. The Geneva Risk and Insurance Review 45, 134–170.
Hong, J., Seog, S.H., 2023. Health insurance system and resilience to epidemics. Risk Anal. 43 (1), 97–114. https://doi.org/10.1111/risa.14005
Kim, S.H., Min, E., Hwang, Y.M., Choi, Y.S., Yi, J.W., 2022a. Impact of COVID-19 pandemic on thyroid surgery in a university hospital in south korea. Cancers (Basel) 14 (17), 4338.
Kim, Y., Gordon, A., Rowerdink, K., Scott, L.H., Chi, W., 2022b. The impact of the COVID-19 pandemic on health care utilization among insured individuals with common chronic conditions. Med Care 60 (9), 673.
Kish, L., Frankel, M.R., 1974. Inference from complex samples. Journal of the Royal Statistical Society. Series B (Methodological) 36 (1), 1–37. Publisher: [Royal Statistical Society, Oxford University Press]. https://www.jstor.org/stable/2984767.
Kotrych, D., Ciechanowicz, D., Pawlik, J., Szyjkowska, M., Kwapisz, B., Madry, M., 2022. Delay in diagnosis and treatment of primary bone tumors during COVID-19 pandemic in poland. Cancers (Basel) 14 (24), 6037.
Leeflang, P., Bijmolt, T., van Doorn, J., Hanssens, D., van Heerde, H., Verhoef, P., Wieringa, J., 2009. Creating lift versus building the base: current trends in marketing dynamics. International Journal of Research in Marketing 26, 13–20. https://doi.org/10.1016/j.ijresmar.2008.06.006
Li, S., Zhuang, J., Shen, S., 2017. Dynamic forecasting conditional probability of bombing attacks based on time-series and intervention analysis. Risk Anal. 37 (7), 1287–1297. https://doi.org/10.1111/risa.12679
Lubitz, J., Prihoda, R., 1984. The use and costs of medicare services in the last 2 years of life. Health Care Financ Rev 5 (3), 117–131.
Michalowsky, B., Hoffmann, W., Bohlken, J., Kostev, K., 2021. Effect of the COVID-19 lockdown on disease recognition and utilisation of healthcare services in the older population in germany: a cross-sectional study. Age Ageing 50 (2), 317–325.

Mizee, M., Schaap, L.A., Hoogendijk, E.O., van Schoor, N.M., 2022. Delay or postponement of medical care among older adults in the netherlands at earlier and later stages of the COVID-19 pandemic. Aging Clin Exp Res 34 (11), 2913–2917.

Mogharab, V., Ostovar, M., Ruszkowski, J., Hussain, S. Z.M., Shrestha, R., Yaqoob, U., Aryanpoor, P., Nikkhoo, A.M., Heidari, P., Jahromi, A.R., et al., 2022. Global burden of the COVID-19 associated patient-related delay in emergency healthcare: a panel of systematic review and meta-analyses. Global Health 18 (1), 58.

Pfeffermann, D., 1993. The role of sampling weights when modeling survey data. International Statistical Review / Revue Internationale de Statistique 61 (2), 317–337. Publisher: [Wiley, International Statistical Institute (ISI)]. https://doi.org/10.2307/1403631

Plott, C.F., Kachalia, A.B., Sharfstein, J.M., 2020. Unexpected health insurance profits and the COVID-19 crisis. JAMA 324 (17), 1713–1714.

Poisal, J.A., Sisko, A.M., Cuckler, G.A., Smith, S.D., Keehan, S.P., Fiore, J.A., Madison, A.J., Rennie, K.E., 2022. National health expenditure projections, 2021–30: growth to moderate as COVID-19 impacts wane: study examines national health expenditure projections, 2021-30 and the impact of declining federal supplemental spending related to the COVID-19 pandemic. Health Aff 41 (4), 474–486.

Przybytniowski, J.W., Borkowski, S., Pawlik, A., Garasyim, P., 2022. The risk of the COVID-19 pandemic and its influence on the business insurance market in the medium- and long-term horizon. Risks 10 (5), 100.

Ricciardiello, L., Ferrari, C., Cameletti, M., Gaianill, F., Buttitta, F., Bazzoli, F., De'Angelis, G.L., Malesci, A., Laghi, L., 2021. Impact of SARS-CoV-2 pandemic on colorectal cancer screening delay: effect on stage shift and increased mortality. Clinical Gastroenterology and Hepatology 19 (7), 1410–1417.

Richter, A., Wilson, T.C., 2020. COVID-19: Implications for insurer risk management and the insurability of pandemic risk. The Geneva Risk and Insurance Review 45, 171–199.

Rodríguez Sánchez, A., Wucherpfennig, J., Rischke, R., Iacus, S.M., et al., 2023. Search-and-rescue in the central mediterranean route does not induce migration: predictive modeling to answer causal queries in migration research. Sci Rep 13 (1), 11014. https://doi.org/10.1038/s41598-023-38119-4

Rosenberg, M.A., Young, V.R., 1999. A bayesian approach to understanding time series data 3 (2), 130–143. https://doi.org/10.1080/10920277.1999.10595808

Schmitt, E., Tull, C., Atwater, P., et al., 2018. Extending bayesian structural time-series estimates of causal impact to many-household conservation initiatives. Ann Appl Stat 12 (4), 2517–2539. Publisher: Institute of Mathematical Statistics. https://doi.org/10.1214/18-AOAS1166

Scitovsky, A.A., 2005. "The high cost of dying": what do the data show? Milbank Q 83 (4), 825. Publisher: Milbank Memorial Fund. https://doi.org/10.1111/j.1468-0009.2005.00402.x

Scott, S.L., 2024. bsts: Bayesian Structural Time Series. R package version 0.9.10, https://CRAN.R-project.org/package=bsts.

Scott, S.L., Varian, H.R., 2014. Predicting the present with bayesian structural time series. International Journal of Mathematical Modelling and Numerical Optimisation 5 (1).

Shumway, R.H., Stoffer, D.S., 2017. ARIMA Models. In: Shumway, R.H., Stoffer, D.S. (Eds.), Time Series Analysis and Its Applications: With R Examples. Springer International Publishing. Springer Texts in Statistics, pp. 75–163. https://doi.org/10.1007/978-3-319-52452-8_3

Szczygielski, J.J., Charteris, A., Bwanya, P.R., Brzeszczyński, J., 2022. The impact and role of COVID-19 uncertainty: a global industry analysis. International Review of Financial Analysis 80, 101837.

Team, R.C., 2019. R: A Language and Environment for Statistical Computing. Place: Vienna, Austria, https://www.r-project.org/.

Wilensky, G.R., 2022. The COVID-19 pandemic and the us health care workforce. JAMA Health Forum 3 (1), e220001–e220001.

Wüthrich, M.V., Merz, M., 2023. Statistical Foundations of Actuarial Learning and its Applications. Springer Nature.

Xu, S., Glenn, S., Sy, L., Qian, L., Hong, V., Ryan, D.S., Jacobsen, S., 2021. Impact of the COVID-19 pandemic on health care utilization in a large integrated health care system: retrospective cohort study. J. Med. Internet Res. 23 (4), e26558.