

Kurz, Christoph F. et al.

**Article — Published Version**

## The effect of bariatric surgery on health care costs: A synthetic control approach using Bayesian structural time series

Health Economics

**Provided in Cooperation with:**

John Wiley & Sons

*Suggested Citation:* Kurz, Christoph F. et al. (2019) : The effect of bariatric surgery on health care costs: A synthetic control approach using Bayesian structural time series, Health Economics, ISSN 1099-1050, Wiley, Hoboken, NJ, Vol. 28, Iss. 11, pp. 1293-1307, <https://doi.org/10.1002/hec.3941>

This Version is available at:

<https://hdl.handle.net/10419/230048>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<http://creativecommons.org/licenses/by-nc/4.0/>

# The effect of bariatric surgery on health care costs: A synthetic control approach using Bayesian structural time series

Christoph F. Kurz<sup>1,2</sup>  | Martin Rehm<sup>1</sup> | Rolf Holle<sup>1</sup> | Christina Teuner<sup>1,3</sup> | Michael Laxy<sup>1,2</sup> | Larissa Schwarzkopf<sup>1</sup>

<sup>1</sup>Institute of Health Economics and Health Care Management, Helmholtz Zentrum München, Neuherberg, Germany

<sup>2</sup>German Center for Diabetes Research, Neuherberg, Germany

<sup>3</sup>Munich School of Management and Munich Center of Health Sciences, Ludwig-Maximilians-Universität München, Munich, Germany

## Correspondence

Christoph F. Kurz, Institute of Health Economics and Health Care Management, Helmholtz Zentrum München Deutsches Forschungszentrum für Gesundheit und Umwelt (GmbH), Ingolstädter Landstraße 1, 85764 Neuherberg, Germany.  
Email: christoph.kurz@helmholtz-muenchen.de

## Abstract

Surgical measures to combat obesity are very effective in terms of weight loss, recovery from diabetes, and improvement in cardiovascular risk factors. However, previous studies found both positive and negative results regarding the effect of bariatric surgery on health care utilization. Using claims data from the largest health insurance provider in Germany, we estimated the causal effect of bariatric surgery on health care costs in a time period ranging from 2 years before to 3 years after bariatric intervention. Owing to the absence of a control group, we employed a Bayesian structural forecasting model to construct a synthetic control. We observed a decrease in medication and physician expenditures after bariatric surgery, whereas hospital expenditures increased in the post-intervention period. Overall, we found a slight increase in total costs after bariatric surgery, but our estimates include a high degree of uncertainty.

## KEYWORDS

bariatric surgery, Bayesian structural model, obesity, synthetic controls, time series

## 1 | INTRODUCTION

Overweight and obesity are a global public health concern. If current trends continue, by 2025, global obesity prevalence will reach 18% in men and surpass 21% in women (NCD Risk Factor Collaboration, 2016). In Germany, the prevalence of overweight has not risen in the last few years, but the prevalence of obesity has continued to increase. Approximately one in five 18- to 79-year-olds has a body mass index of at least 30 kg/m<sup>2</sup> and is therefore considered obese (Mensink et al., 2013). A number of comorbidities are associated with overweight and obesity. People with obesity are at increased risk of type II diabetes mellitus, cardiovascular diseases, and certain types of cancer (Garg, Maurer, Reed, & Selagamsetty, 2014). Ultimately, these conditions lead to a shortened life expectancy compared with people of normal weight (Hu et al., 2004).

Conservative treatment methods include dietary changes, physical activity, and behavioral therapy (Berg et al., 2014). However, surgical intervention is more effective than conservative treatment with respect to reduction in body fat, improvement of obesity-associated diseases, and lowering mortality (Colquitt, Pickett, Loveman, & Frampton, 2014; Jakobsen et al., 2018; Sjöström et al., 2009; Wirth, Wabitsch, & Hauner, 2014). Different types of bariatric surgery are available and are used to counteract existing obesity in very obese individuals. Bariatric surgery may be considered for

individuals with a body mass index greater than or equal to 35 kg/m<sup>2</sup>, when the use of conventional methods has been unsuccessful. (Consensus Development Conference Panel, 1991) Surgical procedures enable substantial and long-term weight loss (Sjöström et al., 2004), leading to a reduction in the incidence of diabetes (Carlsson et al., 2012), cardiovascular events (Sjöström et al., 2012), and neoplasms (Narbro et al., 2002).

On account of the high economic cost of overweight and obesity for health care providers (Tsai et al., 2011; Yates et al., 2016), the economic aspects of bariatric surgery also need to be considered. In Germany, absorption of the costs of bariatric surgical procedures is currently not included in the standard benefit catalog of statutory health insurance (SHI). However, the interventions can be requested for individual cases and financed by the patient's insurance fund. This could be the reason why the number of bariatric interventions in relation to the population is considerably lower in Germany than in neighboring countries. The frequency of operations for morbid obesity is currently 9 per 100,000 adults in Germany; in other European countries, it is many times higher (Sweden: 77; France: 57; and Belgium: 108; Angrisani et al., 2015).

Given the increasing number of obese people and thus increasing costs to SHI funds, an evaluation of the economic impact of bariatric surgery is of great importance. Bariatric surgery is associated with a reduction in the use of medication for the treatment of obesity-related diseases (Keating, Peeters, Swinburn, Carter, & Moodie, 2013; Makary et al., 2010; Narbro et al., 2002; Potteiger et al., 2004), but a reduction in the use of medication alone is not a sufficient proof that a procedure is economically reasonable.

The clinical effects of bariatric surgery are well documented (Wirth et al., 2014), but it is a very expensive procedure, and the question to whether bariatric surgery is cost-effective (i.e., more effective but more costly than usual care) or even cost saving (i.e., more effective and less costly than usual care) is still unanswered (Maciejewski & Arterburn, 2013).

Only a few studies have compared the pre-surgical and post-surgical costs among individuals who underwent bariatric surgery in the main service categories of a health care system. Although three studies showed higher total health expenditures after surgery (Bleich et al., 2012; Maciejewski et al., 2010; McEwen et al., 2010), cost reductions were observed in another (Crémieux et al., 2008). These differing results do not produce a conclusive picture of how a bariatric intervention affects the use of medical services, especially as these studies did not analyze the reasons for utilization. Only one study investigated the economic consequences of bariatric surgery under the conditions of the German health care system (Augurzy et al., 2016). This study used a matched control group to evaluate the effects of bariatric surgery, which can be problematic as relevant inclusion and matching criteria may not be available in the data or not known and even a rich set of covariates is insufficient to control for preexisting differences between the treatment and control groups (Arceneaux, Gerber, & Green, 2006).

As an alternative to estimate the causal impact that bariatric surgery has on health care utilization, we use a “synthetic control” approach. The term synthetic control originally referred to the method of Abadie, Diamond, and Hainmueller (2010) but today includes a broader range of similar methods (Xu, 2017). These methods have been shown to be useful in the analysis of intervention effects through time series data in the absence of a randomized controlled trial. In health economics and health policy evaluation, the synthetic control method (SCM) has been used, for example, to estimate the effect of user fee removal in Zambia (Lépine, Lagarde, & Le Nestour, 2018), pay for performance (Kreif et al., 2016; Ryan, Krinsky, Kontopantelis, & Doran, 2016), and universal health insurance coverage (Courtemanche & Zapata, 2014). For more applications in health research, consult the review by Bottell, Craig, Lewsey, Robinson, and Popham (2018). In this paper, we use a Bayesian structural approach that estimates the model on the pretreatment period using Gibbs sampling and then iterate each sampling trajectory forward using the estimated parameters to construct the post-intervention counterfactual. This counterfactual forecasting method has the advantage that it does not require a set of dedicated control units, that is, untreated individuals, and instead can use any sort of related time series to predict the counterfactual.

The rest of the paper is organized as follows. First, we provide an overview on SCMs and detail the Bayesian structural time-series (BSTS) model in Section 2. In Section 3, we give an overview of the data. We present the results of our analysis in Section 4. Finally, we discuss the results in Section 5 with the implications of the model results and the limitations, before we conclude in Section 6.

## 2 | SYNTHETIC CONTROL METHODS

In this section, we briefly review existing methods for synthetic controls and explain the difference from the Bayesian structural model. We start by defining the general framework and notation, closely following Imbens and Rubin (2015).

In the synthetic control framework, we assume  $T$  time points and  $J + 1$  units, where the first unit,  $j = 0$ , is exposed to the intervention in periods  $T_0 + 1$  to  $T$  and unexposed in periods 1 to  $T_0$ . The synthetic control unit is created from a collection of the remaining  $J$  control units, the so-called donor pool. These units are unexposed in all time periods.

We define  $y_{jt}^1$  as the outcome that would be observed for unit  $j$  at time  $t$  if unit  $j$  is exposed to the intervention in periods  $T_0 + 1$  to  $T$  and  $y_{jt}^0$  as the outcome that would be observed for unit  $j$  at time  $t$  in the absence of the intervention.

The observed outcome  $y_{jt}$  can be written in terms of the potential outcomes framework (Neyman, 1923; Rubin, 1974):

$$y_{0t} = \begin{cases} y_{0t}^0 & \text{for } t = 1, \dots, T_0 \\ y_{0t}^1 & \text{for } t = T_0 + 1, \dots, T \end{cases}.$$

The treatment effect for unit  $j$  at time  $t$  is defined as  $\tau_{jt} = y_{jt}^1 - y_{jt}^0$  for  $t = T_0 + 1, \dots, T$ . This is the treatment effect on the treated. However, the counterfactual  $y_{0t}^0$  is never observed for the post-intervention period. In consequence, to estimate this treatment effect, the missing counterfactual observations have to be imputed. In the present case, the response variable is health care utilization (in particular, we look at medication costs, inpatient costs, and outpatient costs as proxies for utilization), so the causal effect of interest is the difference between the observed series, and the series that would have been observed had the intervention—in our case, bariatric surgery—not taken place. Note that for all pre-intervention periods  $t \in \{1, \dots, T_0\}$ , we have  $y_{jt}^1 = y_{jt}^0$ .

To estimate a treatment effect, it is necessary to make an assumption regarding the outcomes that would have occurred in the absence of treatment for the treated unit  $y_{0t}^0$ . Let  $d_{jt}$  be an indicator variable equal to 1 if unit  $j$  is treated at time  $t$  and 0 otherwise. Furthermore, let  $\mathbf{x}_{jt}$  be observed covariates. We now assume that the treatment-free potential outcome for both groups is the same in expectation conditional on past outcomes and observed covariates:

$$y_{jh}^0 \perp d_{jt} | (\mathbf{y}_j, \mathbf{x}_{jt}),$$

where the symbol  $\perp$  defines statistical independence,  $h > T_0$  marks the post-intervention period, and  $\mathbf{y}_j = (y_{j1}, \dots, y_{jT_0})$  is the set of pre-treatment outcomes. This assumption is called the conditional independence assumption. Under this assumption, individuals with similar outcomes in the pre-treatment period would be anticipated to have similar potential treatment-free outcomes in post-treatment periods after conditioning on observed covariates  $\mathbf{x}_{jt}$  (O'Neill et al., 2016). However, the validity of a particular identifying assumption can never be tested empirically because the counterfactual outcome cannot be observed. In the following, we always assume conditional independence.

## 2.1 | Traditional synthetic control methods

The SCM by Abadie and Gardeazabal (2003) and Abadie et al. (2010) is an alternative approach for evaluating the treatment effect in observational settings without randomized controls that goes beyond simple pre-post comparisons of means. With this method, a number of time series that are unaffected by the intervention are optimally weighted according to their fit to the outcome in the pre-intervention period. They are then combined into a composite time series to which the treatment group is compared. This comparison is used to estimate what would have happened to the treatment group if it had not received the treatment and allows variation of observed and unobserved predictors over time. It is important that the control time series themselves must not be affected by the intervention and the relationship between the control time series and the treated time series, as established during the pre-period, must remain stable throughout the post-period. Essentially, the model parameters are estimated on the pre-intervention data and used to calculate the entire post-intervention path. Formally, the goal is to find a vector of weights  $W = (w_1, \dots, w_J)$  that minimizes the distance in pre-treatment outcomes between the treated and the weighted average of the control units. This can be written as the following optimization problem:

$$W^* = \arg \min_{w \in W} \frac{1}{T_0} \sum_{t=1}^{T_0} (y_{0t} - \sum_{j \neq 0} w_j y_{jt})^2.$$

Abadie et al. (2010) restrict the weights to be positive and sum to 1,  $\sum_j w_j = 1, w_j \geq 0$ . This has the appealing property of being a simple weighted combination of the control units that can be easily interpreted. In addition, Athey and Imbens (2017) point out that very often positive weights are only applied to a subset of the control units; therefore, variable selection is performed as only relevant control units are chosen.

The SCM is not a generative model but exploits linear relationships between treated and control units to construct a counterfactual. Abadie et al. (2010) proved that when both unobserved and observed characteristics of the treated unit lie in the convex hull of the characteristics of control units, the SCM estimator is asymptotically unbiased. In cases where the

treated unit lies outside the distribution of the control units, no convex combination of control units can be constructed, leading to severe bias (Abadie et al., 2010; Ferman & Pinto, 2016).

Recent SCM extensions include elastic net (Zou & Hastie, 2005) regularization (Doudchenko & Imbens, 2016), interactive fixed effects (Xu, 2017), or use matrix completion methods to estimate counterfactual (untreated) outcomes in synthetic control settings (Athey, Bayati, Doudchenko, Imbens, & Khosravi, 2017).

The SCM has proven valuable as a means of relaxing the parallel trends assumption required by additive fixed effect models and difference in differences (DiD). If the synthetic control succeeds in balancing the treated unit outcome with a weighted combination of control units in the pre-intervention period, unmeasured time-varying confounders will also be balanced. This permits additional flexibility, because DiD only allows for measured time-varying confounders.

## 2.2 | Bayesian structural time series

In this paper, we use the BSTS approach to construct the synthetic control. The main difference from the traditional synthetic control approach by Abadie et al. (2010) is that it explicitly models the outcome of the treated unit and uses the control units (which can also be other explanatory predictor time series) to capture remaining variation. That includes using information from the post-intervention period for the control units. The BSTS approach thus produces a dynamic forecast conditional on the pre-intervention trends as well as the pre-intervention and post-intervention values of the predictor time series. In addition, it is not restricted to the class of convex combinations of predictor variables and can more flexibly include time-series effects such as trends and seasonality. BSTS was first proposed by Scott and Varian (2013) and extended to the synthetic control setting by Brodersen, Gallusser, Koehler, Remy, and Scott (2015). The basis of BSTS is a local linear trends model in a state-space setting (Durbin & Koopman, 2012) that incorporates a set of predictive covariates in the observation equation. It starts by defining two equations:

$$\begin{aligned} y_{0t} &= \underset{(1 \times 1)}{\mathbf{Z}_t} \underset{(1 \times d)}{\alpha_t} + \underset{(d \times 1)}{\varepsilon_t} & (\text{observation equation}), \\ \alpha_{t+1} &= \underset{(d \times 1)}{\mathbf{T}_t} \underset{(d \times d)}{\alpha_t} + \underset{(d \times q)}{\mathbf{R}_t} \underset{(q \times 1)}{\alpha_t} & (\text{state equation}), \end{aligned}$$

where  $\varepsilon_t \sim \mathcal{N}(0, \sigma^2)$  is the observation noise and  $\alpha_t \sim \mathcal{N}(\mathbf{0}, \mathbf{Q})$  is a  $q$ -dimensional system error with a  $q \times q$  state diffusion matrix  $\mathbf{Q}$ . Both are independent of each other. The vector  $y_{0t}$  contains the observed outcome at time  $t$  for the treated unit, whereas  $\alpha_t$  is the vector of the latent state process at time  $t$ . The values  $\alpha_t$  are unobserved and determined by the system, therefore describing the state.

$\mathbf{Z}_t$  is a  $(1, d)$ -dimensional matrix containing the regressors that predict  $y_{0t}$ ,  $\mathbf{T}_t$  is a transition matrix that describes evolution of the state vector over time, and  $\mathbf{R}_t$  is a control matrix that defines which states vary over time.

The local linear trends model from above can be directly interpreted by its components. If we decompose the model as a sum of trend component  $\mu_t$  and regression component  $\lambda_t$ , we can rewrite it as follows:

$$y_{0t} = \mu_t + \lambda_t + u_t, \quad u_t \sim \mathcal{N}(0, \sigma_u^2).$$

In this sense,  $\alpha_t$  is now the aggregation of these components:

$$\alpha_t = [\mu_t^T, \lambda_t^T]^T,$$

where all state components are modeled independently, yielding an additive contribution to  $y_{0t}$ . The first component,  $\mu_t$ , is a local linear trend and models the underlying level changes over time, that is, expresses the data as a linear function of time:

$$\begin{aligned} \mu_t &= \mu_{t-1} + \delta_{t-1} + v_t, & v_t &\sim \mathcal{N}(0, \sigma_v^2), \\ \delta_t &= \delta_{t-1} + w_t, & w_t &\sim \mathcal{N}(0, \sigma_w^2). \end{aligned}$$

The error term  $v_t$  allows the level of the trend to shift stochastically, whereas  $w_t$  supports the slope of the trend to change over time. The special case that  $\sigma_v^2 = \sigma_w^2 = 0$  results in deterministic trend. The second component,  $\lambda_t$ , is the regression component with static coefficients  $\beta$ ,

$$\lambda_t = \beta x_t,$$

where  $\mathbf{x}_t$  is the pool of available predictor time series. This decomposition of the time series into several stochastic processes captures the dynamics of the time series. The unknown parameters  $\theta$  in this system are the variance terms and the regression coefficients:

$$\theta : \{\sigma_u^2, \sigma_v^2, \sigma_w^2, \beta\},$$

and are estimated by the Kalman filter model (Grewal, 2011). The Kalman filter has the advantage that it has no problem with nonstationarity and requires no differencing or identification stages. Instead, it updates its estimates when new observations arrive and therefore has a Bayesian interpretation. However, the Kalman filter assumes linearity and is based upon noise terms that are normally distributed (Commandeur, Koopman, & Ooms, 2011). This strong assumption is not always met in practice. If it is not met, the Kalman filter is the best linear estimator, but nonlinear estimators may be better.

Models in the state-space setting are very flexible: All autoregressive integrated moving average (ARIMA) time series models, as well as all linear models, are special cases of state-space models. The state vector and the system matrices are very modular, making it possible to incorporate various features, such as seasonality and circular effects, into this model (Harvey, Trimbur, & Van Dijk, 2007). Please note that we omitted these additional components in the description above for simplicity and did not use them in our analysis.

It is important that the outlined structure remains intact throughout the modeling period. If, for example, local trends disappear over time or the static regression coefficients become dynamic, estimation accuracy may suffer (Qiu, Jammalamadaka, & Ning, 2019).

To estimate the set of model parameters  $\theta$ , we define  $\mathbf{y} = (y_{01}, \dots, y_{0T_0})$  as the observed data in the pre-intervention period and let  $\alpha = (\alpha_1, \dots, \alpha_{T_0})$  denote the corresponding state sequence. We then define a prior distribution  $p(\theta)$  on the model parameters and a distribution  $p(\alpha|\theta)$  on the initial state values. This allows us to sample from the posterior  $p(\alpha, \theta|\mathbf{y})$  using Markov chain Monte Carlo (MCMC) through a Gibbs sampler. We can then draw predictions of the counterfactual from  $p(\theta, \alpha|\mathbf{y})$ . We define an inverse gamma prior for the state error variance parameter and a “spike-and-slab” prior for the regression coefficients. This prior is commonly used in Bayesian regression modeling because it performs variable selection by combining a point mass at 0 (the spike) with a weakly informative distribution on the set of non-zero coefficients (the slab). The spike part is a Bernoulli distribution, and the slab part is a weakly informative normal-inverse-gamma distribution. See Brodersen et al. (2015) for more details and George and McCulloch (1997) for a discussion on Bayesian variable selection. This stochastic approximation to inference using MCMC is used because no closed-form posterior exists for this model.

To estimate the causal effect of the bariatric surgery, we use the estimated states and parameters to forecast trajectories of the treated unit  $\hat{y}_{0t}^0$  for the post-intervention time points  $t = T_0 + 1, \dots, T$ . This procedure is repeated many times. Each draw of parameters from the posterior results in slightly different forecasts that are then averaged. Madigan and Raftery (1994) proved that averaging over an ensemble of models is at least as good as the best single model (see Fragoso, Bertoli, & Louzada, 2018, for a review on Bayesian model averaging). We compare the actual response to the forecasted posterior distribution. This allows us to report the tail-area probability, that is, the probability under the calculated posterior that the response is at least as extreme (away from the expected value) as the observed one. Subtracting this predicted time series from the observed response during the post-intervention period yields a semiparametric Bayesian posterior distribution for the causal effect. We report the pointwise causal impact in this paper.

The BSTS model is asymptotically unbiased because, given that the model outlined above is the true data-generating model, the posterior distribution will converge to a point mass on its true value as the number of post-intervention time points goes to infinity.

The advantage of this Bayesian approach is that we can use these samples from the posterior to report statistics such as the average absolute and relative effect caused by the intervention, including their credible intervals (CIs). The CI can be thought of as the region of highest subjective belief within which an unobserved parameter falls (Jaynes & Kempthorne, 1976). We implement BSTS in the R (R Core Team, 2017) programming language with the `CausalImpact` (Brodersen et al., 2015) package.

### 3 | DATA DESCRIPTION

The analysis is based on quarterly insurance billing claims data from AOK, an association of health insurance plans that serves more than 24 million people and thus almost a third of the German population. We included data from all individuals who underwent bariatric surgery–associated interventions between the second quarter of 2007 and the third quarter of 2010 for whom complete information (i.e., no missing values or incomplete insurance status) was available



over the 5-year period from 2006 to 2011. We decided to chose 2006 as the earliest year of observation because, until this year, outpatient diagnoses were not coded with a sufficient degree of quality. Moreover, we assumed that the introduction of a diagnosis-related group (DRG) system in the hospital sector (2005) changed coding strategies. 2011 is the year when the data agreement ended.

Only patients who underwent sleeve gastrectomy (OPS code 54345), gastric bypass (OPS codes 54454 and 54455), or gastric banding (OPS code 5448) were kept in the analytic cohort. Patients with biliopancreatic diversion (OPS codes 54343, 54344, and 54346) were excluded because of the low number of cases undergoing this procedure. All individuals had a primary diagnosis of obesity (International Statistical Classification of Diseases and Related Health Problems, 10th Edition [ICD-10] codes E65–E68), and no diagnosis of neoplasms of digestive organs before or during index hospitalization (ICD-10 codes C15–C26, D00, D01, and D10–D13) to make sure the focus of the intervention was on the treatment of obesity. We ended up with  $N = 2,733$  individuals who received bariatric surgery in our analysis. Table 1 presents an overview of the baseline characteristics. The mean age at the time of surgery was 43.6 years, with 70.3% females. Some 36.5% of all individuals underwent sleeve gastrectomy, 37.9% gastric bypass, and 25.6% gastric banding surgery.

The observation period for each patient was defined with reference to the date of the first bariatric intervention. A patient may have received multiple bariatric interventions in the observation period, for example if a gastric band had to be replaced. Therefore, we determined the quarter of the bariatric intervention (QBI) by the discharge date of the stay in hospital based on the first bariatric surgery. For example, for an individual discharged from hospital on April 24, 2009, QBI would have been q2/2009. We analyzed the 8 quarters before and the 12 quarters after QBI. To exclude utilization that is directly related to the intervention, we omitted the quarter immediately before and after QBI as well as the QBI itself. Because the QBI is different in each individual, the full observation period is not available for all individuals. Figure 1 shows the number of individuals in each quarter. At least 1,024 individuals were observed in each quarter.

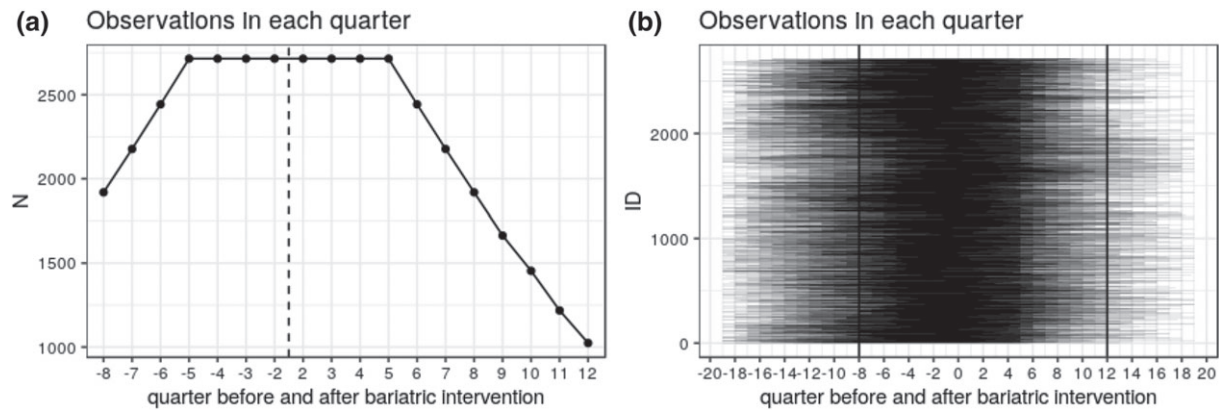
The outcome parameters in our analysis were health care utilization and expenditures in the main service categories of a health care system from the payer perspective: medication, physician care, and hospital care. In our analysis, we aggregated the individual data into a single time series, so the reported outcome is always the average costs per patient. This is a common approach for the analysis of health outcomes and costs in time-series analyses (Getzen, 2006; Lee & Miller, 2002; Xie et al., 2016). Hospital care refers to a defined period of time that may extend over several quarters. Therefore, we assigned the costs of hospitalization proportional to the time of service use within a quarter, assuming constant amounts per day. We standardized cost data to 2016 values using German health care-specific inflation rates to ensure the comparability of costs per quarter.

We also used data related to medication, physician, and hospital costs of bariatric surgeries in Germany. In our BSTS model, we took the total costs of medication, hospital, and outpatient treatments, as well as the costs and prevalences of diabetes and obesity, the prevalence of hypertension, and the unemployment rate in Germany over the years 2006–2011 as covariates (see Figure 2). Because outcome is relative to QBI, we aggregated the control series in the same way as the outcome, that is, calculated control time series for each individual's observation period and then averaged it. All these data

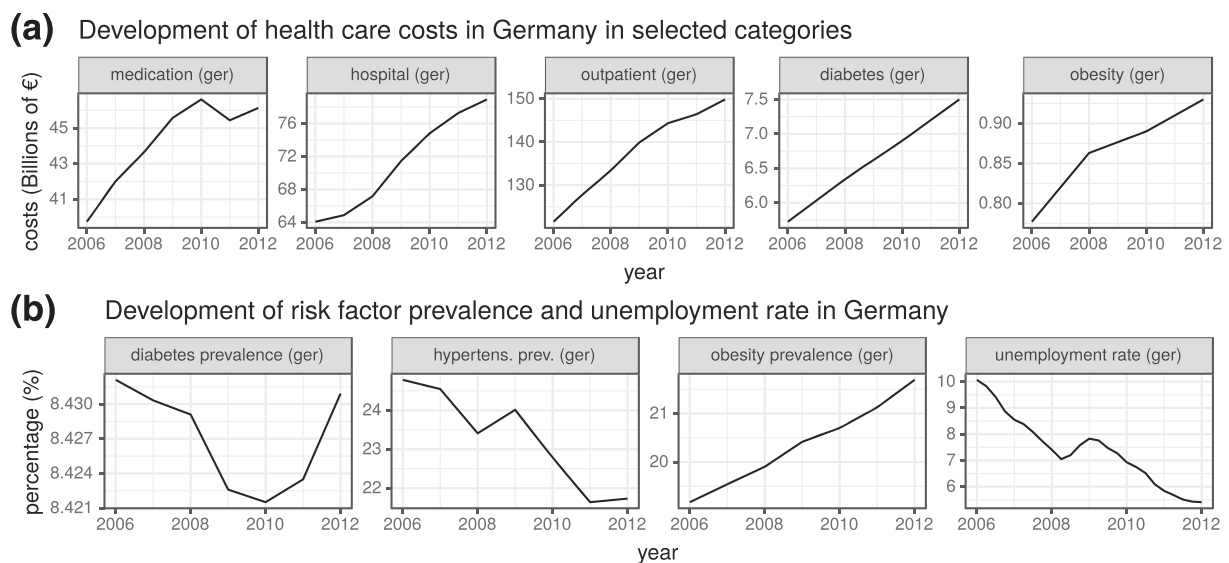
**TABLE 1** Table of baseline characteristics (age, sex, operation type, and selected comorbidities) of all 2,733 individuals at the time of bariatric surgery

	Overall
Age (mean [ <i>SD</i> ])	43.6 (11.0)
Sex, female (%)	1,922 (70.3)
Operation type (%)	
Sleeve gastrectomy	997 (36.5)
Gastric bypass	1,036 (37.9)
Gastric banding	700 (25.6)
Sleep apnea (%)	764 (28.0)
Diabetes (%)	1,314 (48.1)
Arthritis (%)	1,280 (46.8)
Dyslipidemia (%)	1,163 (42.6)
Hypertension (%)	1,987 (72.7)
Observations	2,733

*Note.* Sleep apnea is defined by ICD-10 code G473. Diabetes mellitus is defined by ICD-10 codes E10–E14. Arthritis is defined by ICD-10 codes M15–M19. Dyslipidemia is defined by ICD-10 code E78, and hypertension is defined by ICD-10 codes I10–I15.



**FIGURE 1** (a) Number of observations  $N$  in each quarter before and after bariatric surgery. The dashed vertical line marks the time of the bariatric intervention. (b) Time line for each individual and the observed quarters for each. Vertical lines mark the period used for analysis



**FIGURE 2** Time series used as covariates in the Bayesian structural time series model during the observation period. Upper row (a) shows cost-related covariates, and lower row (b) shows risk factor-related variables and unemployment rate

are freely available from the German Federal Statistical Office ([www.destatis.de](http://www.destatis.de)). Only the costs of diabetes and obesity and the unemployment rate were available quarterly, so we linearly interpolated the in-between quarters for the other covariates. These variables describe the general time trend of health care costs in Germany and specifically for obese patients, and are correlated with the medication, physician, and hospital expenditures in the pre-intervention period in our AOK sample. To ensure that results are invariant to linear transformations of the data, we standardized the covariates. We provide the code for our analysis online (Kurz, 2019).

## 4 | RESULTS

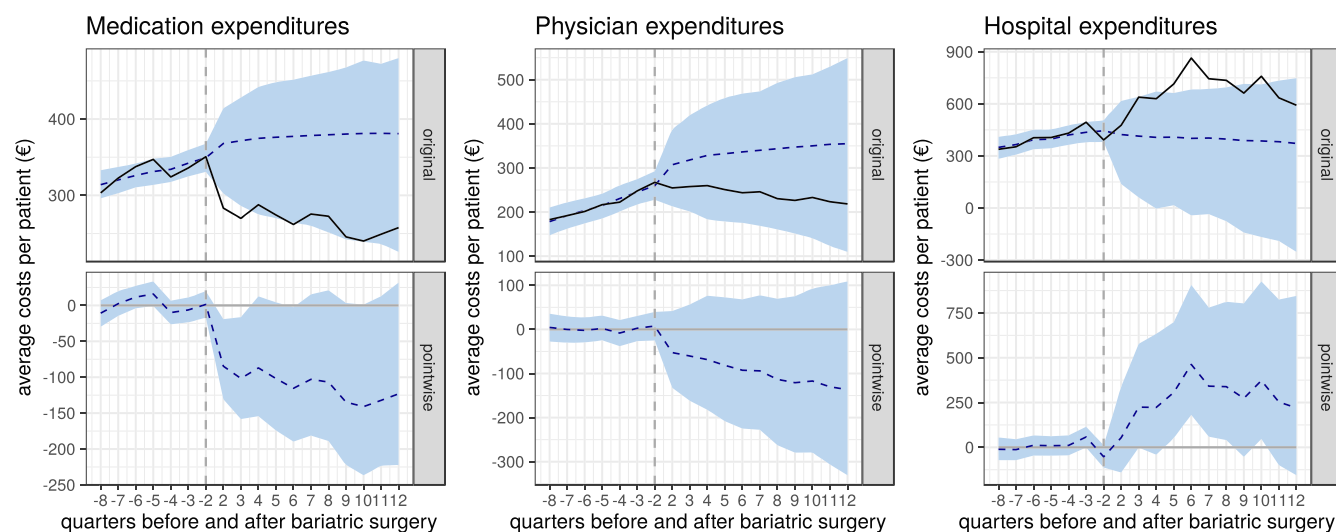
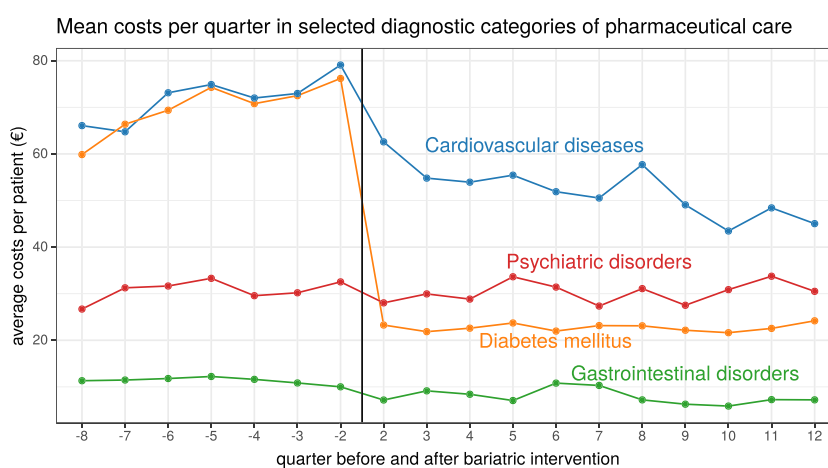
Table 2 shows the results of the BSTS analysis. Figure 3 visualizes the observed and predicted time-series data. All reported values are inflation-adjusted 2016 Euros (€). For medication costs, we find an average value of €265 during the post-intervention period. In contrast, in the absence of an intervention, we would have expected average costs of €373. The 95% CI of this counterfactual prediction is [258, 450]. Subtracting this prediction from the observed cost yields an estimate of the causal effect that the intervention had on the cost variable. This effect is €-108 with a 95% CI of [-183, 4]. In relative terms, the cost variable showed a decrease of 29%. This means that the cost-saving effect observed during the intervention period can be considered statistically significant as the probability of obtaining this effect by chance is small (Bayesian one-sided tail-area  $p = .03$ ). Figure 4 shows the mean costs per quarter in selected diagnostic categories



**TABLE 2** The causal effect of bariatric surgery on health care costs

	Average expenditures after bariatric surgery		
	Medication	Physician	Hospital
Actual	265	240	677
Predicted	373	355	397
95% CI	[258, 450]	[148, 507]	[-59, 677]
Absolute effect	-108	-115	278
95% CI	[-183, 4]	[-266, 92]	[0, 736]
Relative effect	-29%	-32%	70%
95% CI	[-49, 1]	[-75, 26]	[0, 185]
Posterior tail-area <i>p</i>	.03	.07	.02
Posterior <i>p</i> of causal effect	97%	93%	98%

Note. All values are 2016 Euros (€).

**FIGURE 3** Average medication, physician, and hospital expenditures per person and quarter before and after bariatric surgery. The upper plots (“original”) show the observed expenditures (solid black line) and the counterfactual synthetic controls (dashed blue line) including the 95% credible interval according to the Bayesian structural time series model. The lower plots (“pointwise”) show the average difference between the observed and estimated values [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)] **FIGURE 4** Mean costs per quarter in selected diagnostic categories of pharmaceutical care. Cardiovascular diseases are defined by Anatomical Therapeutic Chemical (ATC) codes C01–C10. Diabetes mellitus is defined by ATC code A10. Gastrointestinal disorders are defined by ATC codes A02–A07. Psychiatric disorders are defined by ATC codes N05 and N06. The black vertical line marks the time of the bariatric intervention [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)] 

of pharmaceutical care before and after bariatric intervention. The overall reduction in medication expenditures after bariatric surgery mainly results from cardiovascular diseases and diabetes mellitus: Medication expenditures for these categories are greatly reduced in the post-intervention period. In contrast, medication expenditures for gastrointestinal and psychiatric disorders remain the same after bariatric intervention.

**TABLE 3** Covariate coefficients from the medication, hospital, and physician expenditures, together with the inclusion probabilities from the Bayesian structural time-series model

	Medication			Physician			Hospital		
	Mean	SD	Inclusion <i>p</i>	Mean	SD	Inclusion <i>p</i>	Mean	SD	Inclusion <i>p</i>
Intercept	0.00	0.00	.00	0.00	0.00	.00	0.00	0.00	.00
Medication (ger)	−0.01	2.26	.27	0.65	1.65	.23	2.29	2.28	.49
Hospital (ger)	0.55	1.24	.34	0.47	1.18	.32	−1.47	1.49	.66
Outpatient (ger)	0.38	1.45	.31	0.40	1.37	.30	0.10	1.94	.51
Diabetes (ger)	0.16	1.28	.31	0.29	1.13	.31	−0.46	1.48	.53
Obesity (ger)	0.02	1.37	.32	0.34	1.28	.29	1.86	1.56	.65
Diabetes prevalence (ger)	−0.18	0.29	.50	0.08	0.78	.18	0.73	2.31	.19
Hypertension prevalence (ger)	0.01	0.31	.49	0.08	0.68	.27	0.45	1.56	.31
obesity prev. (ger)	0.01	0.80	.32	0.07	0.75	.31	−0.63	1.52	.31
Unemployment rate (ger)	0.06	0.77	.41	0.05	0.65	.29	−0.20	0.87	.31

We also see a cost-saving effect for physician expenditures. Here, the average value during the post-intervention period is €240. In the absence of the surgery, we would have expected average costs of €355 with a 95% CI of [148, 507]. The causal effect, measured by the difference between the predicted and the observed costs, is €−115 with a 95% CI of [−266, 92]. The relative effect is a decrease of 32%. Because of the Bayesian one-sided tail-area  $p = .07$ , this effect cannot be considered statistically significant.

For hospital expenditures, we see an increase in costs in the post-intervention period. In this period, average costs per patient are €677. If individuals had not received the surgery, we would have expected an average cost of €397. The 95% CI of this counterfactual prediction is [−59, 677]. The causal effect, measured by subtracting this prediction from the observed cost, is €278 with a 95% CI of [0, 736]. The relative increase is 70%. The probability of obtaining this effect by chance is  $p = .02$ . This means that the increasing effect for hospital costs observed during the intervention period is unlikely to result from random fluctuations. In combination, we see an expected total average effect of €55 for medication, physician, and hospital expenditures after bariatric surgery in the post-intervention period.

Table 3 shows the coefficients underlying the predictions from the BSTS model for all three models. The first two columns represent the mean and standard deviation from the distributions of the estimated coefficients from all draws. The third column shows the inclusion probability; that is, the probability that the variable is included in the regression, according to the spike-and-slab prior used for variable selection. Higher absolute values mean a higher effect. In all models, the hospital and outpatient covariates have the highest influence on the model.

In all models, the medication covariate has the largest effect, whereas the hospital covariate is always among the highest inclusion probabilities.

## 5 | DISCUSSION

This study investigated the impact of bariatric surgery on health care costs analyzing a period ranging from 2 years before to 3 years after the intervention. Bariatric surgery was associated with a cost reduction in pharmaceuticals and physician services but also with rising costs for inpatient care. In total, we saw a slight increase in total costs after bariatric surgery.

### 5.1 | Effect of bariatric surgery on costs

The cost reduction in pharmaceuticals after bariatric surgery is attributable to the lower costs of treating diabetes and cardiovascular diseases and is in accordance with several previous studies that investigated pharmaceutical utilization in a pre-surgical and post-surgical comparison (Keating et al., 2013; Makary et al., 2010; Neovius et al., 2012; Potteiger et al., 2004).

Expenditures on hospital care were higher in the post-surgical period but with decreasing trend in the longer term. When examining the reasons for hospital admissions in the post-operative period, the literature finds three main reasons that explain the majority of the cost increase (Weiner et al., 2013). First, particularly during the first post-operative year, the reason for admission is often the treatment of complications presumably related to the intervention, such as incisional hernias or syndromes of the operated stomach. Second, as a result of weight loss after bariatric intervention, further treatments such as plastic surgery to reduce skin tissue or the treatment of gallstone disease are required. Third, some treatments can only be performed after successful weight loss following surgery, such as hip or knee joint replacement.

This is supported by the decreasing trend of hospital expenditures after the sixth post-intervention quarter (i.e., after 1.5 years; see Figure 3).

Our findings showing higher total costs after surgery are in line with previous studies, which found increased expenditures after bariatric intervention (Bleich et al., 2012; Maciejewski et al., 2010; McEwen et al., 2010). Other studies also found decreasing costs for medication and physicians but increased inpatient spending (Weiner et al., 2013). The trend of decreasing inpatient costs over time was confirmed in other studies (Czernichow et al., 2015; Neovius et al., 2012). Contrary to our analysis, most of these studies did not account for the increasing trend of expenditures in the pre-intervention period. Still, because this trend is based on extrapolated data from the pre-period, it is not clear how long this extrapolation may be valid or how strongly influenced it is by outliers.

The results obtained here are in contrast to the findings of some existing studies, which show that significant cost savings can be expected after bariatric surgery (Crémieux et al., 2008; Makary et al., 2010). A study in the United Kingdom found bariatric surgery to be cost-effective, where the incremental cost-effectiveness ratios for surgery ranged between £2,000 and £4,000 per quality adjusted life-year (Picot et al., 2009). A study of bariatric patients in the United States revealed that the initial investment is returned within 4 years of the intervention (Crémieux et al., 2008). The authors attributed the cost reductions after bariatric surgery to reductions of comorbidities and to a lower use of medical services for their treatment. This relation could be seen in our analysis at least for the use of pharmaceuticals and physician care, but these were offset by an increase for hospital care.

## 5.2 | Choice of data source

Our study is one of the first that examines the development of health care expenditures after bariatric surgery using data from Germany. The analyses are based on claims data from an SHI provider. These data have the advantage that real costs are depicted and a large number of individuals can be considered. The only other study under the regularities of the German health care system by Augurzky et al. (2016) is also based on claims data but from a different provider with fewer cases. As in our study, Augurzky et al. (2016) find increased hospital costs, reduced medication costs, and little change in physician costs. They take into account the direct costs of the surgery itself and calculate the causal effect by a matched control group.

We chose the payer perspective for the cost analysis. This perspective disregards out-of-pocket payments, and indirect costs such as productivity losses cannot be taken into account. Physical complaints associated with obesity often result in the incapacity to work. A recent study assessed the functional capacity of patients before and after bariatric surgery in a cross-sectional study (Vargas, Picolli, Dani, Padoin, & Mottin, 2013). Improvements in test outcomes were seen only 3 months after surgery, which allows the presumption that an increase in productivity could be expected.

There are other limitations to this study regarding the data source. First, only expenditures in the three main service areas of SHI were considered in the analyses. About 70% of the total expenditures of the SHI arise in these domains, but spending on rehabilitation services, for example, is not included in this study. Second, the analyses are based on claims data from a single health insurance fund. Because the characteristics of the insured can differ greatly between health insurance companies, the generalizability of the results is restricted (Jaunzeme, Eberhard, & Geyer, 2013). However, the market share of AOK to SHI for members is nearly 35%. In addition, baseline characteristics of individuals in our study are in line with the German Bariatric Surgery Registry, which collects most bariatric surgery procedures in Germany ( $N = 22,364$ , mean age = 42.6, female = 72.4%; Stroh, Weiner, Wolff, Knoll, & Manger, 2014).

## 5.3 | Choice of model

As in many natural experiment settings, we did not have a randomized control group. The provided data set contained only cases, so no statistical matching procedures (e.g., based on propensity scores) were possible. We did not have a set of untreated individuals. Therefore, we used a BSTS approach to construct a synthetic control, based only on covariates that are predictive for health care costs. This model has advantages over several traditional approaches. Although, for example, classical regression models already go beyond DiD schemes, in that they account for full counterfactual trajectories, they disregard serial autocorrelation and do not allow temporal variation of the coefficients (Brodersen et al., 2015).

The BSTS approach has several advantages over the traditional SCM. Many of the SCM limitations arise from its restriction to convex combinations of the control units. Ferman and Pinto (2016) show that the SC estimator is biased if treatment assignment is correlated with unobserved heterogeneity and if treatment assignment is only correlated with time-invariant-unobserved variables. BSTS is not restricted to convex combinations of the control units and has additional variable selection, making it less affected by this type of bias. A simulation study by Kinn (2018) compares SCM and BSTS in terms of mean squared error (MSE) and bias. The author considers both parallel and heterogeneous trend

cases, a case where a convex combination of controls is insufficient to recreate the treated unit, a case where trends are not easily detected because of noise, and a case with a short pretreatment period and high noise. Kinn (2018) finds that BSTS is superior or equal to SCM in all these cases. Especially in cases that are similar to our study (difficult trends and noise, weakly associated control time series), BSTS is the most accurate method for finding the true treatment effect with the lowest MSE, bias, and variance. (Kinn, 2018) also argues that the increased flexibility of methods such as BSTS can better capture the data-generating process than SCM. However, BSTS cannot deal with multiple treated units as the SCM can (Kreif et al., 2016).

Other classes of models, for example ARIMA models, can account for autocorrelation but only among observations and not latent states. This inhibits the ability to distinguish between state noise and observation noise (Leeftang et al., 2009). The fully Bayesian setting further allows incorporation of empirical priors on the parameters and takes model uncertainty into account when forecasting the counterfactual. The spike-and-slab prior allows inclusion of covariates that may not be predictive, as it automatically selects only the relevant ones and avoids several of the disadvantages of other variable selection methods (George & McCulloch, 1997).

There is evidence that medical expenditures are subject to seasonal effects (Rolden, Rohling, van Bodegom, & Westendorp, 2015), but including a stochastic dummy component to incorporate quarterly seasonal effects into the model did not change the results.

Choice of priors had little effect on the results, as comparing them with the posteriors showed that they were overwhelmed by the data.

## 5.4 | Choice of covariate time series

Traditional applications of the SCM construct the counterfactual response by time series based on a comparable group in a different country (Billmeier & Nannicini, 2013) or by having a set of unaffected individuals (Pechlivanoglou, Wieringa, Jager, & Postma, 2015). The increased flexibility of BSTS allows for a much broader choice of covariates. For example, Scott and Varian (2013) included dozens of correlated time series from Google Trends as covariates in their BSTS model.

We used data related to our outcomes as covariables where there might not be a strong causal relationship. A similar approach was chosen by Elleby, Hansen, and Yu (2015) to predict the counterfactual rice and wheat prices in India, by Shioda et al. (2019) to estimate the impact of vaccination, and by de Vocht (2019) for the impact of mobile phone use on brain cancer.

Bayesian structural time-series is, in principle, a forecasting method. It can be used without any covariate time series. For example, a study by Qiu et al. (2019) found that the prediction accuracy of BSTS is better than ARIMA. There are also studies that use no covariate time series to construct a counterfactual. The study by Marschik et al. (2018) constructs the counterfactual only based on trend calculation of the pre-intervention period to estimate the effect of a bovine vaccination program. However, using no covariates should be avoided. A study by Gallego, Angulo, Suárez-García, and Gómez-Ullate (2018) compared three different BSTS models according to their accuracy to estimate the effect of an advertising campaign: one with no covariates, one with lagged outcome as covariate, and one with external covariates. As in our study, the covariates were not a set of untreated control units but covariates with assumed predictive information, for example, price index, unemployment rate, rainfall, and temperature. Gallego et al. (2018) find that these external regressors, even if only weakly associated with the outcome, greatly improve BSTS accuracy and treatment effect estimation.

In general, it is difficult to judge whether a covariate is valid or similar to the outcome or constitutes a good fit (Bouttell et al., 2018). Because all covariables were selected by the model, we consider them appropriate in our analysis. We believe that our linear between-quarters interpolation is adequate because all covariates show a relatively linear trend (see Figure 2).

## 5.5 | Limitations

Our results include a large amount of uncertainty. This means that, although it may look as though bariatric surgery shows a small total increase in costs (€55), considering the intervention period as a whole, the 95% posterior probability intervals for the effects in each of the three categories of health care costs (medication, physician, and hospital expenditures) include 0 in all cases. Therefore, the effect estimates have to be interpreted with caution. They could be the result of random fluctuations that are unrelated to the intervention. According to Brodersen et al. (2015), this can be the case when the intervention period is too short to distinguish signal from noise, when there are not enough control variables, or when the control variables do not correlate well with the response variable during the learning period. Another solution would be to specify more informative prior knowledge into the model, but this could lead to unrealistically small CIs.

In recent years, there have been major changes in the application of surgical procedures in practice. In the considered analysis period, gastric bands were more commonly used than they are currently (Stroh et al., 2016). This is because gastric banding is less costly but also less effective and more often needs revision. Studies suggest that, with regard to diabetes remission, gastric bypass is more effective than gastric banding (Weiner et al., 2014). However, in this evaluation, the cost trends for gastric bypass and gastric banding did not differ when analyzed by type of surgery.

Our analysis only covers a 3-year observation period after surgery, including 1,024 individuals in the last quarter. This makes it difficult to see the long-term effects of bariatric surgery. Studies on the medical effects of bariatric surgery have shown reductions in obesity-related diseases such as diabetes and cardiovascular diseases. Although we already see this effect for medication and physician expenditures, the effect on hospital expenditures may become more apparent in a longer post-intervention period. High costs for the treatment of complications related to bariatric interventions immediately after surgery should decrease over time. This effect is already apparent in the later post-intervention quarters for hospital expenditures in our analysis. In a longer observation period, it should be even more evident.

In our analysis, the number of included individuals varies over quarters, especially in the later post-intervention quarters. However, we are only interested in the average effect of all individuals, and this effect should stay stable because we still have many individuals and there is no underlying selection mechanism as to why an individual is not included. This “censoring” is informative and does not cause bias, because it is unrelated to their disease status or health care utilization (Ranganathan & Pramesh, 2012). Average age, sex, and operation types are nearly constant during all quarters, but it might affect the statistical power of this study. In addition, the reliance on a single set of identification assumptions may not be adequate but is a common practice.

Finally, because we do not include a (randomized) nonsurgical control group, we cannot conclude that changes in costs after bariatric intervention definitely resulted from the procedure.

## 5.6 | Sensitivity analysis

There is no way to check whether all assumptions that the BSTS approach requires are fulfilled (see Section 2), but we performed several sensitivity analyses to verify the reliability and robustness of our results. First, we included only individuals who were observed over the whole period ( $N = 191$ ). In this analysis, the direction of the effect stays the same as in the main analysis, but the cost-saving effect for medication expenditures after bariatric surgery is bigger, whereas the costs for hospital expenditures do not increase as much as in the main analysis (Appendix A.1). However, the small sample size makes the analysis more vulnerable to outliers.

In the second analysis, we bootstrapped the data to have  $N = 5,000$  observations in each quarter. Here, the absolute effects are slightly smaller than in the main analysis, especially for hospital expenditures. Relative effects are very similar to the main analyses (Appendix A.2).

The third analysis does not use any covariate time series. Here, the counterfactual is only based on linear trend calculation of the pre-intervention period. Again, the results are very similar to the main analysis, but the linear trend may be overestimating the effects. This model also has very large confidence bands due to the increased uncertainty (Appendix A.3).

In a fourth sensitivity analysis, we checked the validity of the covariate time series. This model included only random noise time series as covariates, sampled from a standard normal distribution. It is very similar to the previous one; except the hospital expenditures are estimated to be lower (Appendix A.4). This is likely to be because of the random variation of the Gibbs sampler due to the large amount of uncertainty in this model. This sensitivity analysis proves the relevance and predictive importance of our included covariates in the main analysis.

## 6 | CONCLUSION

Despite these limitations, this study presents an important comparison of health care expenditures before and after bariatric surgery and shows temporal trends based on data from Germany. Our findings indicate that, over a time horizon of 3 years after bariatric surgery, medication costs and physician costs are substantially reduced, but the costs of inpatient care increase. Our estimates include a large amount of uncertainty due to the study limitations through the absence of a control group. Routine data are an important source for further research, but larger samples and longer follow-up are needed. For a comprehensive economic evaluation, patient-reported data on indirect costs and on health outcomes have to be analyzed together with direct medical costs.



## ACKNOWLEDGEMENTS

The authors thank the AOK Research Institute (WIdO) for providing the data for this study. Moreover, they are grateful for helpful comments from Martin Wabitsch and Christine Stroh on an earlier version of this manuscript. Remarks from the associate editor, Richard Grieve, and two anonymous reviewers greatly improved the manuscript.

## ORCID

Christoph F. Kurz  <https://orcid.org/0000-0001-9498-8002>

## REFERENCES

- Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *J. Am. Stat. Assoc.*, 105(490), 493–505.
- Abadie, A., & Gardeazabal, J. (2003). The economic costs of conflict: A case study of the Basque country. *Am. Econ. Rev.*, 93(1), 113–132.
- Angrisani, L., Santonicola, A., Iovino, P., Formisano, G., Buchwald, H., & Scopinaro, N. (2015). Bariatric surgery worldwide 2013. *Obesity Surgery*, 25(10), 1822–1832.
- Arceneaux, K., Gerber, A. S., & Green, D. P. (2006). Comparing experimental and matching methods using a large-scale voter mobilization experiment. *Political Analysis*, 14(1), 37–62.
- Athey, S., Bayati, M., Doudchenko, N., Imbens, G., & Khosravi, K. (2017). Matrix completion methods for causal panel data models. *arXiv preprint arXiv:1710.10251*.
- Athey, S., & Imbens, G. W. (2017). The state of applied econometrics: Causality and policy evaluation. *The Journal of Economic Perspectives*, 31(2), 3–32.
- Augurzky, B., Wübker, A., Pilny, A., Fels, K., Bohm, S., Priess, H. W., ..., & Lehmann, B. (2016). Barmer GEK Report Krankenhaus 2016: Barmer GEK.
- Berg, A., Bischoff, S. C., Colombo-Benkmann, M., Ellrott, T., Hauner, H., Heintze, C., ..., & Teufel, M. (2014). Interdisziplinäre leitlinie der qualität s3 zur prävention und therapie der adipositas, *Deutsche Adipositasgesellschaft, Deutsche Diabetes Gesellschaft, Deutsche Gesellschaft für Ernährung, Deutsche Gesellschaft für Ernährungsmedizin*.
- Billmeier, A., & Nannicini, T. (2013). Assessing economic liberalization episodes: A synthetic control approach. *Review of Economics and Statistics*, 95(3), 983–1001.
- Bleich, S. N., Chang, H.-Y., Lau, B., Steele, K., Clark, J. M., Richards, T., ..., & Segal, J. B. (2012). Impact of bariatric surgery on healthcare utilization and costs among patients with diabetes. *Medical Care*, 50(1), 58.
- Bouttell, Janet, Craig, Peter, Lewsey, James, Robinson, Mark, & Popham, Frank (2018). Synthetic control methodology as a tool for evaluating population-level health interventions. *Journal of Epidemiology and Community Health*, 2000–2017.
- Brodersen, K. H., Gallusser, F., Koehler, J., Remy, N., & Scott, S. L. (2015). Inferring causal impact using Bayesian structural time-series models. *Annals of Applied Statistics*, 9(1), 247–274.
- Carlsson, L. M. S., Peltonen, M., Ahlin, S., Anveden, Å., Bouchard, C., Carlsson, B., ..., & Näslund, I. (2012). Bariatric surgery and prevention of type 2 diabetes in Swedish obese subjects. *The New England Journal of Medicine*, 367(8), 695–704.
- Colquitt, J. L., Pickett, K., Loveman, E., & Frampton, G. K. (2014). Surgery for weight loss in adults. *Cochrane Libr.*
- Commandeur, J. J. F., Koopman, S. J., & Ooms, M. (2011). Statistical software for state space methods. *Journal of Statistical Software*, 41(i01).
- Consensus Development Conference Panel (1991). Gastrointestinal surgery for severe obesity. *Annals of Internal Medicine*, 115(12), 956–961.
- Courtemanche, C. J., & Zapata, D. (2014). Does universal coverage improve health? The Massachusetts experience. *Journal of Policy Analysis and Management*, 33(1), 36–69.
- Crémieux, P.-Y., Buchwald, H., Shikora, S. A., Ghosh, A., Yang, H. E., & Buessing, M. (2008). A study on the economic impact of bariatric surgery. *The American Journal of Managed Care*, 14(9), 589–96.
- Czernichow, S., Moszkowicz, D., Szwarcensztein, K., Emery, C., Lafuma, A., Gourmelen, J., & Fagnani, F. (2015). Impact of bariatric surgery on the medical management and costs of obese patients in France: An analysis of a national representative claims database. *Obesity Surgery*, 25(6), 986–996.
- de Vocht, F. (2019). Analyses of temporal and spatial patterns of glioblastoma multiforme and other brain cancer subtypes in relation to mobile phones using synthetic counterfactuals. *Environmental Research*, 168, 329–335.
- Doudchenko, Nikolay, & Imbens, Guido W (2016). Balancing, regression, difference-in-differences and synthetic control methods: A synthesis: National Bureau of Economic Research.
- Durbin, J., & Koopman, S. J. (2012). *Time series analysis by state space methods*, Vol. 38: OUP Oxford.
- Elleby, C., Hansen, H., & Yu, W. (2015). Domestic price and welfare effects of the 2007–11 Indian grain export restrictions. In *Agricultural and Applied Economics Association & Western Agricultural, 2015 aaea & waea joint annual meeting, july 26-28, san francisco, california*.
- Ferman, B., & Pinto, C. (2016). Revisiting the synthetic control estimator.
- Fragoso, T. M., Bertoli, W., & Louzada, F. (2018). Bayesian model averaging: A systematic review and conceptual classification. *International Statistical Review*, 86(1), 1–28.
- Gallego, V., Angulo, P., Suárez-García, P., & Gómez-Ullate, D. (2018). Assessing the effect of advertising expenditures upon sales: A Bayesian structural time series model. *arXiv preprint arXiv:1801.03050*.

- Garg, S. K., Maurer, H., Reed, K., & Selagamsetty, R. (2014). Diabetes and cancer: Two diseases with obesity as a common risk factor. *Diabetes, Obesity and Metabolism*, 16(2), 97–110.
- George, E. I., & McCulloch, R. E. (1997). Approaches for Bayesian variable selection. *Statistica Sinica*, 339–373.
- Getzen, T. E. (2006). Aggregation and the measurement of health care costs. *Health Services Research*, 41(5), 1938–1954.
- Grewal, M. S. (2011). Kalman filtering. *International encyclopedia of statistical science*: Springer, pp. 705–708.
- Harvey, A. C., Trimbur, T. M., & Van Dijk, H. K. (2007). Trends and cycles in economic time series: A Bayesian approach. *Journal of Economics*, 140(2), 618–649.
- Hu, F. B., Willett, W. C., Li, T., Stampfer, M. J., Colditz, G. A., & Manson, J. E. (2004). Adiposity as compared with physical activity in predicting mortality among women. *The New England Journal of Medicine*, 351(26), 2694–2703.
- Imbens, G. W., & Rubin, D. B. (2015). *Causal inference in statistics, social, and biomedical sciences*: Cambridge University Press.
- Jakobsen, G. S., Småstuen, M. C., Sandbu, R., Nordstrand, N., Hofso, D., ..., & Hjelmæsæth, J. (2018). Association of bariatric surgery vs medical obesity treatment with long-term medical complications and obesity-related comorbidities. *Journal of the American Medical Association*, 319(3), 291–301.
- Jaunzeme, J., Eberhard, S., & Geyer, S. (2013). How representative are SHI (statutory health insurance) data? *Bundesgesundheitsblatt*, 56(3), 447–454.
- Jaynes, Edwin T., & Kempthorne, Oscar (1976). Confidence intervals vs Bayesian intervals, *Foundations of probability theory, statistical inference, and statistical theories of science*: Springer, pp. 175–257.
- Keating, C. L., Peeters, A., Swinburn, B. A., Carter, R., & Moodie, M. L. (2013). Pharmaceutical utilisation and costs before and after bariatric surgery. *International Journal of Obesity*, 37(11), 1467–1472.
- Kinn, D. (2018). Synthetic control methods and big data. *arXiv preprint arXiv:1803.00096*.
- Kreif, N., Grieve, R., Hangartner, D., Turner, A. J., Nikolova, S., & Sutton, M. (2016). Examination of the synthetic control method for evaluating health policies with multiple treated units. *Health Economics*, 25(12), 1514–1528.
- Kurz, C (2019). Bach project repository. <https://github.com/krz/BACH>
- Lee, R., & Miller, T. (2002). An approach to forecasting health expenditures, with application to the US medicare system. *Health Services Research*, 37(5), 1365–1386.
- Leeflang, P. S. H., Bijmolt, T. H. A., Van Doorn, J., Hanssens, D. M., Van Heerde, H. J., Verhoef, P. C., & Wieringa, J. E. (2009). Creating lift versus building the base: Current trends in marketing dynamics. *International Journal of Research in Marketing*, 26(1), 13–20.
- Lépine, A., Lagarde, M., & Le Nestour, A. (2018). How effective and fair is user fee removal? Evidence from Zambia using a pooled synthetic control. *Health Economics*, 27(3), 493–508.
- Maciejewski, M. L., & Arterburn, D. E. (2013). Cost-effectiveness of bariatric surgery. *Journal of the American Medical Association*, 310(7), 742–743.
- Maciejewski, M. L., Smith, V. A., Livingston, E. H., Kavee, A. L., Kahwati, L. C., Henderson, W. G., & Arterburn, D. E. (2010). Health care utilization and expenditure changes associated with bariatric surgery. *Medical Care*, 48(11), 989–998.
- Madigan, D., & Raftery, A. E. (1994). Model selection and accounting for model uncertainty in graphical models using Occam's window. *Journal of the American Statistical Association*, 89(428), 1535–1546.
- Makary, M. A., Clark, J. M., Shore, A. D., Magnuson, T. H., Richards, T., Bass, E. B., ..., & Segal, Jodi B (2010). Medication utilization and annual health care costs in patients with type 2 diabetes mellitus before and after bariatric surgery. *Archives of Surgery*, 145(8), 726–731.
- Marschik, T., Obritzhauser, W., Wagner, P., Richter, V., Mayerhofer, M., Egger-Danner, C., ..., & Pinior, B. (2018). A cost-benefit analysis and the potential trade effects of the bovine viral diarrhoea eradication programme in Styria, Austria. *Veterinary Journal*, 231, 19–29.
- McEwen, L. N., Coelho, R. B., Baumann, L. M., Bilik, D., Nota-Kirby, B., & Herman, W. H. (2010). The cost, quality of life impact, and cost-utility of bariatric surgery in a managed care population. *Obesity Surgery*, 20(7), 919–928.
- Mensink, G. B. M., Schienkiewitz, A., Haftenberger, M., Lampert, T., Ziese, T., & Scheidt-Nave, C. (2013). Übergewicht und adipositas in deutschland: Ergebnisse der studie zur gesundheit erwachsener in deutschland (degsl1)(leitthema). *Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz*, 56(5-6), 786–794.
- NCD Risk Factor Collaboration (2016). Trends in adult body-mass index in 200 countries from 1975 to 2014: A pooled analysis of 1698 population-based measurement studies with 19.2 million participants. *The Lancet*, 387(10026), 1377–1396.
- Narbro, K., Ågren, G., Jonsson, E., Näslund, I., Sjöström, L., & Peltonen, M. (2002). Pharmaceutical costs in obese individuals: Comparison with a randomly selected population sample and long-term changes after conventional and surgical treatment: The SOS intervention study. *Archives of Internal Medicine*, 162(18), 2061–2069.
- Neovius, M., Narbro, K., Keating, C., Peltonen, M., Sjöholm, K., Ågren, G., ..., & Carlsson, L. (2012). Health care use during 20 years following bariatric surgery. *Journal of the American Medical Association*, 308(11), 1132–1141.
- Neyman, Jerzy (1923). On the application of probability theory to agricultural experiments. Essay on principles. Section 9. *Statistical Science*, 465–472.
- O'Neill, S., Kreif, N., Grieve, R., Sutton, M., & Sekhon, J. S. (2016). Estimating causal effects: Considering three alternatives to difference-in-differences estimation. *Health Services & Outcomes Research Methodology*, 16(1-2), 1–21.
- Pechlivanoglou, P., Wieringa, J. E., Jager, T., & Postma, M. J (2015). The effect of financial and educational incentives on rational prescribing. A state-space approach. *Health Economics*, 24(4), 439–453.

- Picot, J., Jones, J., Colquitt, J. L., Gospodarevskaya, E., Loveman, E., Baxter, L., & Clegg, A. J. (2009). The clinical effectiveness and cost-effectiveness of bariatric (weight loss) surgery for obesity: A systematic review and economic evaluation. *Health Technology Assessment*, 13(41), 1–190.
- Potteiger, C. E., Paragi, P. R., Inverso, N. A., Still, C., Reed, M. J., Strodel, W., ..., & Petrick, A. (2004). Bariatric surgery: Shedding the monetary weight of prescription costs in the managed care arena. *Obesity Surgery*, 14(6), 725–730.
- Qiu, J., Jammalamadaka, S. R., & Ning, Ning (2019). Multivariate Bayesian structural time series model. *Journal of Machine Learning Research*, 19(68).
- R Core Team (2017). R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org>
- Ranganathan, P., & Pramesh, C. S. (2012). Censoring in survival analysis: Potential for bias. *Perspectives in Clinical Research*, 3(1), 40.
- Rolden, H. J. A., Rohling, J. H. T., van Bodegom, D., & Westendorp, R. G. J. (2015). Seasonal variation in mortality, medical care expenditure and institutionalization in older people: Evidence from a Dutch cohort of older health insurance clients. *PLoS ONE*, 10(11), e0143154.
- Rubin, Donald B (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Education & Psychology*, 66(5), 688.
- Ryan, A. M., Krinsky, S., Kontopantelis, E., & Doran, T. (2016). Long-term evidence for the effect of pay-for-performance in primary care on mortality in the UK: A population study. *The Lancet*, 388(10041), 268–274.
- Scott, S. L., & Varian, H. R. (2013). Bayesian variable selection for nowcasting economic time series: National Bureau of Economic Research.
- Shioda, Kayoko, Schuck-Paim, Cynthia, Taylor, Robert J, Lustig, Roger, Simonsen, Lone, Warren, Joshua L, & Weinberger, Daniel M (2019). Challenges in estimating the impact of vaccination with sparse data. *Epidemiology*, 30(1), 61–68.
- Sjöström, Lars, Gummesson, Anders, Sjöström, CDavid, Narbro, Kristina, Peltonen, Markku, Wedel, Hans, ..., & Dahlgren, Sven (2009). Effects of bariatric surgery on cancer incidence in obese patients in Sweden (Swedish obese subjects study): A prospective, controlled intervention trial. *The Lancet Oncology*, 10(7), 653–662.
- Sjöström, Lars, Lindroos, Anna-Karin, Peltonen, Markku, Torgerson, Jarl, Bouchard, Claude, Carlsson, Björn, ..., & Sjöström, C. D. (2004). Lifestyle, diabetes, and cardiovascular risk factors 10 years after bariatric surgery. *The New England Journal of Medicine*, 351(26), 2683–2693.
- Sjöström, L., Peltonen, M., Jacobson, P., Sjöström, C. D., Karason, K., Wedel, H., ..., & Bergmark, G. (2012). Bariatric surgery and long-term cardiovascular events. *Journal of the American Medical Association*, 307(1), 56–65.
- Stroh, C., Köckerling, F., Volker, L., Frank, B., Stefanie, W., Christian, K., ..., & Thomas, M. (2016). Results of more than 11,800 sleeve gastrectomies. *Annals of Surgery*, 263(5), 949–955.
- Stroh, C., Weiner, R., Wolff, S., Knoll, C., & Manger, T. (2014). Are there gender-specific aspects in obesity and metabolic surgery? Data analysis from the German Bariatric Surgery Registry. *Visceral Medicine*, 30(2), 125–132.
- Tsai, A. G., Williamson, D. F., & Glick, H. A. (2011). Direct medical cost of overweight and obesity in the USA: A quantitative systematic review. *Obesity Reviews*, 12(1), 50–61.
- Vargas, C. B., Picolli, F., Dani, C., Padoin, A. V., & Mottin, C. C. (2013). Functioning of obese individuals in pre- and postoperative periods of bariatric surgery. *Obesity Reviews*, 23(10), 1590.
- Weiner, R., El-Sayes, I., Manger, T., Weiner, S., Lippert, H., Stroh, C., & Group, Obesity Surgery Working (2014). Antidiabetic efficacy of obesity surgery in Germany: A quality assurance nationwide survey. *Surgery for Obesity and Related Diseases*, 10(2), 322–327.
- Weiner, J. P., Goodwin, S. M., Chang, H.-Y., Bolen, S. D., Richards, T. M., Johns, R. A., ..., & Clark, J. M. (2013). Impact of bariatric surgery on health care costs of obese persons: A 6-year follow-up of surgical and comparison cohorts using health plan data. *JAMA Surgery*, 148(6), 555–561.
- Wirth, A., Wabitsch, M., & Hauner, H. (2014). The prevention and treatment of obesity. *Deutsches Ärzteblatt International*, 111(42), 705.
- Xie, Y., Schreier, G., Hoy, M., Liu, Y., Neubauer, S., Chang, D. C. W., ..., & Lovell, Nigel H (2016). Analyzing health insurance claims on different timescales to predict days in hospital. *Journal of Biomedical Informatics*, 60, 187–196.
- Xu, Y. (2017). Generalized synthetic control method: Causal inference with interactive fixed effects models. *Political Analysis*, 25(1), 57–76.
- Yates, N., Teuner, C. M., Hunger, M., Holle, R., Stark, R., Laxy, M., ..., & Wolfenstetter, S. B. (2016). The economic burden of obesity in Germany: Results from the population-based KORA studies. *Obesity Facts*, 9(6), 397–409.
- Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society, Series B*, 67(2), 301–320.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**How to cite this article:** Kurz CF, Rehm M, Holle R, Teuner C, Laxy M, Schwarzkopf L. The effect of bariatric surgery on health care costs: A synthetic control approach using Bayesian structural time series. *Health Economics*. 2019;28:1293–1307. <https://doi.org/10.1002/hec.3941>