RUHR
ECONOMIC PAPERS

Boris Augurzky
Arndt R. Reichert
Christoph M. Schmidt

The Effect of a Bonus Program for Preventive Health Behavior on Health Expenditures

RUHR UNIVERSITÄT BOCHUM RUB rwi

#373

Imprint

Ruhr Economic Papers

Published by

Ruhr-Universität Bochum (RUB), Department of Economics
Universitätsstr. 150, 44801 Bochum, Germany

Technische Universität Dortmund, Department of Economic and Social Sciences
Vogelpothsweg 87, 44227 Dortmund, Germany

Universität Duisburg-Essen, Department of Economics
Universitätsstr. 12, 45117 Essen, Germany

Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI)
Hohenzollernstr. 1-3, 45128 Essen, Germany

Editors

Editorial Office

Ruhr Economic Papers #373

Responsible Editor: Christoph M. Schmidt

ISSN 1864-4872 (online) – ISBN 978-3-86788-428-0

Ruhr Economic Papers #373

Boris Augurzky, Arndt R. Reichert, and Christoph M. Schmidt

The Effect of a Bonus Program for Preventive Health Behavior on Health Expenditures

Bibliografische Informationen der Deutschen Nationalbibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über: *http://dnb.d-nb.de* abrufbar.

http://dx.doi.org/10.4419/86788428
ISSN 1864-4872 (online)
ISBN 978-3-86788-428-0

Boris Augurzky, Arndt R. Reichert, and Christoph M. Schmidt[1]

The Effect of a Bonus Program for Preventive Health Behavior on Health Expenditures

Abstract

This paper contributes to the analysis of policy measures that attempt to reduce health care expenditures of insurers. We examine the impact of a cash bonus program for preventive health behavior of a German health insurer on prevention effort and health care expenditures using a unique administrative dataset that covers all insurants of the health insurer between 2003 and 2008. We find that the program has been successful in both increasing individual prevention effort and achieving net savings every year since its implementation in 2004. However, while the estimated effect on health care expenditures is statistically significant in the first year, the effects for the second, third, and fourth years turn insignificant. In the fifth year, results for net savings are sensitive in terms of statistical significance when accounting for dynamic selection into the treatment.

JEL Classification: I10, I12, I18

Keywords: Financial incentives; health care expenditures; dynamic treatment effect; health prevention; inverse probability weighting; bonus program

October 2012

---

1 Boris Augurzky, RWI; Arndt R. Reichert, RWI; Christoph M. Schmidt, RWI, Ruhr-Universität Bochum, IZA Bonn, CEPR London. *– The authors are grateful to Anne Delges, Heiko Friedel, Michael Friedrichs, Alfredo R. Paloyo, Alfons Schrör, and to participants of the 2010 annual conference of the Scottish Economic Society as well as the BeNA Research Seminar on Labor Research for their valuable comments..* *– All correspondence to Arndt R. Reichert, RWI, Hohenzollernstr. 1-3, 45128 Essen, Germany, E-Mail: arndt.reichert@rwi-essen.*

1 Introduction

For decades, health expenditures in the majority of industrialized countries have been rising with growth rates higher than GDP growth rates, challenging the sustainability of financing the health care system. Effective policy measures to stabilize health care expenditures have been insufficiently studied. An obvious and popularly suggested measure is to improve public health by increasing prevention efforts. For instance, the World Health Organization has demanded that governments create an economic, social, and institutional environment that promotes risk prevention (WHO 2002 and IUHP 1999).

Several countries have adopted national risk-prevention strategies and have implemented health care reforms that support healthy living (Weinbrenner et al. 2007). Germany, for instance, implemented the Statutory Health Insurance Modernization Act in 2004 which allowed insurers to award bonuses to their clients who participate in preventive measures. Today, the majority of German health insurers offer such bonus programs.

The present paper evaluates a cash bonus program for health prevention of a large public health insurance company in Germany. By contrasting the estimated program savings with its costs, we aim to answer the question of whether health prevention actually results in reduced health care expenditures as it is often suggested. From an economic perspective, the promotion of health-conscious behavior is justified if individuals fail to achieve optimal levels of health prevention. For instance, individuals with non-constant discounting or patients who fear the results of medical tests have been shown to engage insufficiently in preventive measures (Byrne and Thompson 2001, Wu 2003). Moreover, individuals may underinvest in preventive measures because some of the benefits do not accrue to them. An example are vaccinations which generate positive externalities (Giuffrida and Gravelle 1998).

The empirical literature that documents the effectiveness of financial incentives to increase participation in preventive measures is extensive. In a survey, Achat et al. (1999) find that incentives effectively increase immunization uptake. In a randomized trial, Slater et al. (2005) find similar results for preventive-screening engagement. Charness and Gneezy (2009) go one step further and perform a randomized experiment to test the effectiveness of financial bonuses in raising gym attendance and their impact on allied health indicators. They conclude that bonus payments increase total physical activity, yielding health improvements. Similarly, based on data retrieved from a randomized experiment, Augurzky et al. (2012) establish that financial incentives contingent on weight loss are effective.

Cash bonus programs for health prevention have been previously evaluated. Stock et al. (2010) carry out an empirical cost-benefit analysis of another sickness fund's bonus program for preventive health behavior.[1] They find that the health insurer saved €177 per year for insurants who participated in check-ups, prevention courses, immunization, and exercises during three years. Taking program costs into account, net savings amounted to approximately €100 per year and participant. Friedrichs et al. (2009) find average benefits of €130 per participating insured for the bonus programs of 74 German company health insurance funds.

We improve upon these previous works in two important ways. First, we rely on data which provide a much more precise picture of the insured's morbidity. Second, we examine whether our results are robust with respect to a dynamic extension of the econometric model, revisiting the findings of Stock et al. (2010) and Friedrichs et al. (2009). Since individuals may select into the bonus program during several years, the outcome in one year possibly determines subsequent program participation decisions. This could affect the estimates of the benefits of participating during more than one year in the program. Our analysis examines the general importance of accounting for dynamic selection patterns in applied causal treatment evaluation studies, contributing to novel developments in causal analysis.

[1] In particular, Stock et al. (2010) evaluate the bonus program of BARMER (now BARMER-GEK).

The remainder of the paper is organized as follows. Section 2 describes the bonus program under study and the data. In Section 3, we present the estimation strategy. The empirical results are reported in Section 4, and Section 5 concludes.

2 The bonus program and data

The prevention bonus program under study is designed in order to meet the legal requirements stated in § 65a of the German Social Code Book V. The program offers four types of interventions: check-ups and screenings (three interventions), immunization (one intervention), primary prevention (four interventions), and exercise (two interventions). Table 1 gives an overview. Check-ups and screening interventions are composed of check-ups for children, prenatal care, male and female adult screenings for risk factors and onset of chronic diseases, including cancer screenings for both men and women. Predominantly, screenings address cardiovascular diseases, diabetes, and breast (women) and prostate (men) cancers. The immunization intervention consists of a full vaccine protection where the physician verifies whether participants are protected against a pre-defined set of diseases. If participants lack immunization against a certain illness, the physician advises them to get vaccinated; otherwise, they fail the immunization criterion. Similarly, the physician confirms the non-smoking status and the normal-weight status after verifying that participants have not smoked during the last six months and that they have a current body mass index (BMI) between 21 and 27 points. Exercise interventions consist of participating in accredited exercise classes, having a membership in sports clubs or fitness studios, or earning the German sports badge.[2]

— Table 1 —

[2]The German sports badge is a decoration of the German Olympic Sports Federation approving the completion of certain requirements for swimming, jumping power, speed, physical strength, and endurance.

Any member who successfully engages in three obligatory and in two elective measures receives a reduction of the contribution rate of €40 per year. While all check-ups and screenings constitute obligatory measures, the remaining six interventions are elective. Interventions are generally required annually except the male and female adult check-ups, which are mandatory only every second year. Men and women who engaged in these check-ups during the previous year fulfill the criteria automatically. Check-ups, screenings, the basic set of vaccinations, the confirmation of the non-smoking status, and the BMI are covered by the health insurer. Licensed exercise and prevention courses are usually co-financed by the health insurer. The cost for the German sports badge is marginal.

The prevention bonus program was advertised by the sickness fund in newsletters for members. Detailed information and application forms, the so-called "bonus book", were sent on request. All insurants above the age of 16 can opt into the program. Bonus payments are awarded only if the bonus book contains stamps given by the providers of the distinct prevention measures confirming successful participation. Once a year, participants have to hand in their bonus book. If not, they are presumed to have quit the program.

In 2004, 1,011 insured individuals joined the bonus program, of whom 603 continued in 2005 (Table 2).[3] Finally, 357 insured continually participated throughout the period 2004-2008. The total number of participants per year increased constantly, except for 2006, when the number rose steeply (probably due to a rise of the bonus from €40 to €50 in January).

— Table 2 —

Our unique panel dataset derives from administrative sources. It covers all insurants of the health insurer; it includes a wide range of information at the individual level over the period 2003–2008. The data provide very extensive information about yearly expenses for hospital care, prescription drugs, work incapacity, and other services which include, for example, therapeutic

[3]For the sample exclusion restrictions and the sample description, see Section 2.

and rehabilitation measures. Information about the expenses on prevention courses is particularly helpful because it allows examining whether the program increases individual prevention effort. If prevention effort actually rises, we can expect the program to affect health expenditures and the health status of participants. Moreover, the data provide information on several morbidity indicators, such as the primary International Classification of Diseases (ICD)-10 diagnosis of work incapacity, the ICD-10 hospital discharge diagnosis, and the Anatomical Therapeutic Chemical (ATC) code of prescribed drugs. While the data on work incapacity and hospital diagnoses are used to construct dummy variables on the three-digit level, we generate a count variable with the number of drug prescriptions for each main ATC category. All hospital information is underreported for the years 2003 and 2004 because hospitals were obliged to provide information about hospital treatments only starting from July 2004. But, in 2004, many were not yet prepared to do so.[4]

We also observe whether insurants successfully participated in the bonus program in a particular year, i.e., whether they qualified for the bonus payment. Furthermore, the dataset provides information on personal characteristics. However, only information on the family structure of the insured, such as the number of children or having a partner, is available on a yearly basis. All other personal characteristics are only known at the time when the bonus program was launched. The variables are sex, age, living in East Germany[5], educational level[6], the type of occupation,[7] which indicates the required skills at work, and the sector of occupation.[8] Unfortunately, we have no information on the income level, but we assume that education, the type of occupation, and the sector of occupation jointly proxy for the income level. Nevertheless, the income level should

[4]For more details about the reporting obligation, see the German Social Code Book V § 301.

[5]Frequently, this variable bases on the place of employment instead of the residence.

[6]No professional education, professional education, senior high school, technical college and university.

[7]Unskilled blue-collar workers, skilled blue-collar workers, white-collar workers, employees with part time jobs (mini- and midi-jobs), and trainees. Technicians and the so-called “homeworkers” are assigned into the group of skilled blue-collar and white-collar workers, respectively.

[8]Manufacturing, metallic, engineering, social service sector, administration services, and the remaining services.

not significantly affect the demand for health services because co-payments are low.[9]

We exclude all insurants from the analysis who exit or enter the insurer during 2003 and 2008, i.e., if they are not insured with the health insurer at least 350 days in every period. For reasons of comparability, we standardize the cost data to present the average over 365 days. After these exclusions, 58,618 individuals and 351,708 person–time observations remain.

Table 3 gives an overview of the descriptive statistics. In the sample, there are 3.5 percent unemployed individuals, 8.1 percent pensioners and 14.5 percent co-insured. The majority has professional training alone (40 percent), around 3 percent hold a technical-college degree, and 5 percent a university degree. The majority of the insured are white-collar workers (30 percent), while 20 percent are blue-collar workers. The insured have, on average, 0.11 co-insured children, which seems to be very low. When considering only insurants above the age of 18 with a partner, the average number of co-insured children rises to 0.86.

The table further displays the mean for different health expenditures per insurants and year. Hospital expenses are the most substantial (€ 487). On average, they amount to approximately 48 percent of total health expenditures. Expenses on pharmaceuticals (€ 224) are the second-largest cost category (27 percent in 2005 and 32 percent in 2008). Work inability payments comprise around 14 percent of total health expenditures. In Germany, employers pay the full salary during the first six weeks of sick leave. Afterwards, the health insurance covers 70 percent of gross income. With around € 1.9 per insured expenses for prevention courses were relatively low. However, since 2005, they have increased constantly by roughly € 1 per year, i.e., 86 percent p.a. on average.

[9]For the co-insured, i.e., family members who do not have to pay the insurance premium, the unemployed, and the pensioners, we neither observe their educational background nor their occupational skills. In order not to lose these observations, we include indicator variables for being co-insured, unemployed, and pensioner. Data on either education or the sector of occupation is missing for a few observations. Some are lacking information on the sector of occupation as well. Thus, we include three additional dummy variables: one indicating insured with missing information on education, one for missing information on the type of occupation, and a third one for missing information on education, the type of occupation and the sector of occupation. Also for the unemployed, for instance, all three dummy variables ("educational information missing", "sectoral information missing" and "educational and occupational information missing") take on the value 1.

— Table 3 —

Unfortunately, the dataset does not provide any information about expenditures for outpatient services, in particular doctor visits.[10] Thus, we do not observe whether there are any substitution effects from inpatient to outpatient services due to the bonus program (or vice versa).[11] Nevertheless, substitution effects should not show up in the balances of the health insurer. This is because, until 2008, total remuneration was paid in advance for the following year as a per head capitation fee. Thus, the whole budget was fixed and in the short term, costs associated with doctor visits should not be affected by the bonus program.

If program participation were assigned randomly in every period, individual characteristics and pre-treatment outcomes of participants and non-participants would most likely be balanced. In this study, however, program participation depends on individual choice. It is probable that different individuals make different enrollment choices. For illustrative purposes, we describe the differences between participants and nonparticipants in their individual characteristics in Table 4. Because individuals can have different participation states during 2004 and 2008 depending on their enrollment decisions, we compare those individuals who entered the program in the first year with those who did not, regardless of their later participation choices. We focus on the pre-treatment values of the whole set of covariates including the pre-treatment health care expenditures in different areas of health. These variables are potential correlates of both the participation decision and the post-treatment health outcomes under study.

— Table 4 —

[10]Expenditures for outpatient medical care averaged €38.4 billion in 2007, representing 15.2 percent of total German health care expenditures (KBV 2009).

[11]One may think of a program participant with a curable cancer who takes part in the obligatory screening measures of the bonus program. Once cancer is detected, the next step is usually a surgery. The same insurant in the alternative scenario of non-participation would possibly show up at the general practitioner with certain health complaints some months later and maybe many more times until the cancer is finally diagnosed. Thus, taking part in the screening determines in which cost categories health expenses come from.

Columns (1) to (4) of Table 4 present the mean of the personal characteristics of program participants and non-participants in 2004. Program participants have higher mean expenditures for pharmaceuticals and medical utilities. They also seem to have attended prevention courses more often prior to the program launch. This is in line with expectations because these individuals benefit from windfall gains and are simultaneously more likely to invest in their health. Moreover, program participants are significantly older, have better education, and work in jobs which require higher skills. Among the participants, there are more women, more insured from East Germany, and the shares of unemployed and co-insured are smaller. Participants are also more likely to work in the service sector.

3 Estimation strategy

We briefly describe the methodological approach pursued in this paper which is a combination of a difference-in-differences (DiD) and an inverse-probability-weighting (IPW) estimator with regression adjustment. The goal is to identify the average treatment effect on the treated (ATET), i.e., the average effect of participating in the bonus program on the health expenditures of program participants. The ATET is of primary interest here because it represents the target group of such prevention programs.[12] It is the difference between the observable outcome of the program participants and their unobservable counterfactual outcome had they not participated.

Computing a credible estimate of the counterfactual outcome is the main evaluation problem. A simple approach is to take the outcome of a non-participant. However, this is only valid if selection into the bonus program is random, which is not the case here. Therefore, it is necessary to account for the potential selection bias. In this paper, we use the method of propensity-score weighting (Rosenbaum and Rubin 1983, Hirano and Imbens 2001). It constructs a pseudo-

[12] Average treatment effects, on the other hand, provide information about the saving potential when all insured individuals take part in the program rather than the subgroup of insured who self-select into the program.

population by weighting individuals with the inverse of their estimated probability of program participation. Then, this so-called IPW estimator is simply the difference of the weighted means of the outcomes of participants and non-participants.

The IPW estimator is based on the assumption of unconfoundedness of the treatment assignment with the potential outcomes conditional on the propensity score. This assumption requires all variables that affect the likelihood of receiving the treatment to be observed and accounted for in the estimation of the propensity scores. Since the dataset provides extensive information on personal characteristics and on health conditions, we are confident that we are able to control for the relevant factors that affect the participation decision.[13] If, conditional on these variables, selection into treatment can be considered as a random event, the difference between outcomes of the treatment and the comparison groups is a consistent estimate of the ATET.

Robins and Rotnitzky (1995) show that estimation yields consistent results as long as the conditional mean of the outcomes given the covariates is correctly specified (see also Rubin, 1979; Joffe et al., 2004). This opens an alternative way of dealing with the problem of selection into treatment which consists of including all factors that influence the outcome as covariates in the regression of the outcome variable on the treatment indicator (without weighting). We refer to this approach as regression adjustment. As suggested by Hirano and Imbens (2001), a third way is the combination of IPW estimation with regression adjustment, where the weights are estimated based only on those covariates that are relevant determinants of the enrollment decision.[14] Apart from the treatment indicator, all covariates that have a considerable influence on the outcome variable are included in the regression adjustment. This procedure increases the common support without sacrificing consistency. Thus, it allows us to include more observations in the analysis which become scarce in subsequent investigations.

[13] Since hospital information for the years 2003 is underreported, the number of hospital visits, the length of stay in the hospital, the total spending on hospitals, and any three-digit ICD-10 hospital discharge diagnosis are not considered as covariates.

[14] Augurzky and Schmidt (2001) show that including only those covariates in the propensity score estimation that are necessary in terms of their impact on the selection into treatment is a convenient strategy to increase the common support.

More specifically, following the same approach described by Hirano and Imbens (2001), we include only those covariates into the propensity-score estimation which have a significant slope coefficient at the 5-percent level in each bivariate regression of the treatment indicator and the covariates. The regression adjustment is based on covariates with a significant slope coefficient at the 5-percent level in the weighted bivariate regression with the outcome variable. As covariates for the regression adjustment, only pre-treatment variables are included.[15,16] Due to the under-reporting of pre-treatment hospital data, we ex-ante exclude this information from the regression adjustment.

Hirano and Imbens (2001) point out that this selection rule is easy to implement but not necessarily optimal. For example, it does not take into account correlations between the different covariates. But alternative options require an estimation of the basic models for all subsets of regressors which is prohibitively expensive in terms of computing time. The advantage of this procedure is the reduction of small sample as well as asymptotic bias of the IPW estimator (cf., Abadie and Imbens, 2006) increasing the robustness of the estimator.

In addition to controlling for selection into treatment, we apply the DiD estimator to calculate the ATET. More precisely, we calculate the difference of the outcome variable Y in period t with the outcome variable one year before the launch of the program (period 0): $Y_{i,t} - Y_{i,0}$. The DiD estimator eliminates potential effects of time-invariant unobservables on the outcome variable. The remaining bias can originate from only two sources: the considered groups may respond differently to changes in the health care market and macroeconomic conditions but this is not likely the case.

The panel structure of the dataset allows us to estimate the effect of participating in the bonus

[15] Hirano and Imbens (2001) only include those covariates into the estimation of the propensity score which are significantly correlated with the treatment indicator. They observe a similar criterion for the regression adjustment. As selection rule for the covariates, they look at the strength of the marginal correlation between the treatment and each of the covariates separately in the selection equation, and on the conditional correlation of the outcome and each of the covariates given the treatment in the outcome equation. More precisely, they test the null hypothesis that the slope coefficient of each bivariate regression is equal to zero.

[16] In line with Hirano and Imbens (2001), besides the covariates X of insured i, we include an interaction term of the treatment indicator P with the covariates subtracted by their sample average $\left((X_i - \overline{X}_i) \times P_i\right)$.

program for more than one year. While in period 0, the bonus program has not yet started, in period 1, the insured have two options: participation P and non-participation N. The same applies to all subsequent periods. This means that in the second period, four possible combinations could occur: PP, PN, NP, NN. In the fifth year, there are 32 combinations, with insurants who consistently participate during five consecutive years belonging to sequence $PPPPP$ and those who never participate to sequence $NNNNN$.

We are interested in the effect of participating during up to five consecutive years in the program because engaging in prevention only once might hardly affect health outcomes, while consistent program participation could produce the intended benefits. To clarify thoughts, consider the effect of participating in period 2 (but not in period 1) on health expenditures in period 2 which we denote $\theta_2^{NP;NN}$ and the effect of participating in both periods 1 and 2 on health expenditures in period 2, which we consistently denote $\theta_2^{PP;NN}$. If the effectiveness of prevention increases with the number of periods of continuous preventive engagement then $\theta_2^{PP;NN}$ exceeds $\theta_2^{NP;NN}$ in absolute terms. To calculate $\theta_2^{PP;NN}$, we estimate the probability of participating in both years $\Pr(PP)$ which is inverted and used as weight for individuals of sequence PP in the regression. Analogously, the weight for individuals of sequence NN is $1/\Pr(NN)$. The weights are normalized to sum up to one. The difference of the weighted means of the change in the outcome variable of individuals with sequence PP and those with NN all in the common support, i.e., the intersection of the distributions of propensity scores for the participants and the non-participants, yields $\theta_2^{PP;NN}$.[17] A detailed description of the stepwise estimation procedure is described in the Appendix.

For statistical inference, we compute standard errors by employing an estimated variance that

[17]In order to achieve comparability across the effects, all effects are estimated in a way that they apply to first-year participants. Thus, a more complete formal notation would be $\theta_2^{PP;NN}(P)$ and $\theta_2^{NP;NN}(P)$, where the majuscule in parenthesis indicates to which treatment group the effects apply. Please note that the effects $\theta_2^{PP;NN}(P)$ and $\theta_2^{PP;NN}(N)$ apply to two distinct groups of insured who are most likely different in several characteristics, wherefore, both effects most likely differ. The establishment of the ATET requires that all weights are multiplied by $\Pr(P)$. Thus, the weight for individuals of sequence NN, in fact, changes to $\Pr(P)/\Pr(NN)$ in the example above.

exploits the fact that the mean potential outcomes are weighted means of the observed outcomes in the respective sequences. Moreover, by estimating the variance and expectation conditional on the weights, we explicitly acknowledge the randomness of the weights. This variance estimator is suggested by Lechner (2009) and is based on the nearest-neighbor method. In line with Lechner (2009), the number of neighbors is set to $2\sqrt{N}$.

4 Results

4.1 Estimation of the program effects

In the last two columns of Table 4, we present results of the probit estimation for program participation in the first year of the program, applying the control-variable selection rule. The results confirm the above discussion about which individual characteristics determine enrollment in the bonus program. All significant ICD-10 work inability diagnoses, among those dystonia and hypertensive heart disease with (congestive) heart failure, are positively correlated with program participation, indicating that enrollment is not predominantly an issue of healthy insurants. Figure 1 shows the distribution of the estimated propensity scores for program participants and non-participants of 2004.[18] The distribution for the non-participants is skewed to the left with a very high density at propensity scores close to zero. On average, although program participants have a higher propensity score than non-participants, it is still low. Thus, there is a substantial overlap between both groups.

— Figure 1 —

[18] Although we did not define ex-ante reference categories due to the selection rule, for each of the category groups (type of education, type of occupation, and sector of occupation), the reference groups consist of insurants with characteristics that belong to one of the categories which are left out of the probit estimation. For instance, insurants that have a professional training or a senior high school degree only constitute the reference category for the type of education in the propensity score estimation of program participation in 2004 (cp. Table 4).

Propensity-score weighting without regression adjustment already produces credible estimates of the causal effects of treatment if balancing of the observable covariates is achieved. Columns (5) to (8) of Table 4 show that weighting with the inverse of the estimated propensity scores achieves a good balance of the covariates except for the number of inability cases. However, if this variable is considerably correlated with any of the outcome variables, it will be corrected by the regression adjustment, thus, reducing a potential bias in $\theta_1^{P;N}$.

Results from the DiD-IPW estimation of the effects on both total health care expenditures and on the expenses for prevention courses for up to five years of participation are presented in Table 5. The expenses for prevention courses serve as proxies for changes in individual prevention effort. In the first year of participation, participants spend, on average, €3.94 more on prevention courses than non-participants. This effect is significant and of substantial magnitude compared to the sample average spending of €2 in that year. While the effect on expenses for prevention courses decreases in the second year of participation, its size slightly rises again in the third year. In the fourth and fifth year, the effects are highest. Note that the actual prevention effort may be even larger because some participants take part in prevention interventions which do not show up in this cost category, e.g., the sports club memberships or the full vaccine protection.

— Table 5 —

Concerning total health care expenditures, we find a significant decrease of around €185 for the first year of participation.[19] In the second year, the mean reduction in total expenditures is smaller and no longer statistically significant. In the third year, we observe a larger effect in absolute terms as compared to the two previous years which is insignificant. A continuous participation for four years leads to a statistically insignificant reduction of €308 on average. In the fifth year, we observe a statistically significant (at the 10 percent level) effect of €524. For comparison, yearly total mean health expenditures range from €900 to €1000 in the period 2003–

[19] The estimate for $\theta_2^{NP;NN}$ is also negative and significant giving support to this result.

2008. Thus, the effects are economically meaningful.

To check whether the results are mainly driven by few observations with very large weights, we remove observations with the highest-five and lowest-five percent weights following the trimming procedure of Lechner (2009). We find that the results are robust to this change (Table 5).[20]

An extensive explanation of the above finding of a strong program effect in the first year of participation is beyond the scope of this paper. However, by way of three examples, we illustrate why these short-term effects are reasonable. First, in 2004, 48 percent of the participants engaged in exercise interventions. Charness and Gneezy (2009) find that a few weeks of physical activity improves health indicators within five months. This similarly applies to individuals with certain diseases, e.g., arthritis (Callahan et al., 2008). Second, we estimate that, among the 77 percent of participants who fulfilled the vaccination criterion, 20 percent received a flu vaccination. The latter are participants who have a high risk of flu infection, e.g., elderly people. In contrast to all other insurants, they do not have to co-pay for a flu shot and are, at the same time, required to vaccinate against influenza in order to pass the vaccination criterion. Because a high cost burden is attributable to influenza in Germany (Szucs et al., 2001), and flu shots reduce the risk of hospitalization by 27 percent among elderly people every season (Nichol et al., 2007), this intervention is very likely to be cost-saving in the short-term. Finally, 71 percent of the participants engaged in medical screenings which test patients for diabetes, for example. Patients who were detected to be at risk of diabetes have been shown to prevent or delay the development of diabetes and diabetes-related complications (Engelgau et al., 1998; Chatterjee et al., 2010).[21] Moreover, diabetes screening has been shown to significantly reduce health care costs within at most three years (Chatterjee et al., 2010).

[20] This also holds when we only remove observations with the highest-one and lowest-one percent weights, and for all other results reported in this paper.

[21] In the UK, 60 percent of diabetes diagnoses are made before patients reported any symptom.

A complementary interpretation is anchored on the fact that the participants with the highest potential savings are precisely those people for which health-conscious behavior would yield the highest benefits. These are individuals who are pre-disposed to have or already suffer from illnesses that can be mitigated by certain preventive health interventions. For example, individuals who suffer from lumbago or low back pain would stand to benefit from an exercise program that relieves their back pain and thus eliminates the need for further medication. In this case, the prevention is the cure.

On the other hand, the result of no program effect for two more years of participation might be explained by adult check-ups revealing certain diseases which require treatment likely incurring large health expenditures. Although some diseases may require immediate treatment, medical treatment of most diseases occurs with a certain time lag because the detection is usually followed by a phase of observation of the disease, of diagnostic clarification and, finally, of waiting for the treatment appointment.

The large point estimates for the effects of the fourth and fifth participation year, in contrast, may be due to dynamic selection effects. For instance, the result of cancer screening may indicate the presence of cancer. In this case, participants are unlikely to continue with the program because they would subsequently be under medical treatment. Since these participants actually exit the program, they are excluded from the regression. Such non-random program exits may create a selection problem possibly yielding overestimated treatment effects. This applies, for instance, if participants who exit the program have incurred particularly large health care expenditures which generally applies to cancer patients. Investigating the magnitude of the potential bias is the aim of the subsequent section.

4.2 Assessing the potential bias of the static causal model

An intuitive solution to the problem of non-random program exits might be to control for variables realized in the previous year. This approach, however, triggers the problem of endogeneity because post-program outcomes are influenced by the participation. To circumvent this, Lechner and Miquel (2010) suggest an explicit dynamic causal framework based on the pioneering work of Robins (1986) (see also Lechner, 2009). This framework allows assessing the causal effects of different treatment sequences (e.g., four years of continuous participation) by explicitly taking into account the timing of treatment choices. Variables that jointly influence outcome and selection at each step of the sequential selection process are controlled for. In place of estimating the likelihood of an insured to belong to a certain treatment sequence as in the static causal model, the participation likelihoods are now estimated at each node of a sequence conditional on previous treatment states, where the control and outcome variables enter the regression equation with their past values.

The inverse of the product of these participation likelihood estimates are used as weights in the IPW estimation. Hence, the only difference to the static model lies in the use of other weights. These dynamic weights are largest for those participants who are least likely to enter the program in the first year and most likely to quit in subsequent years. Accordingly, they ensure that those non-participants with the highest likelihood of entering the program in each year have the highest weight among insurants of the comparison group, i.e., the (dynamic) IPW estimation compares two groups of insured who could equally belong to the comparison sequence. Lechner (2009) shows that this dynamic estimator yields consistent estimates of the treatment effect.

Since dynamic selection effects may, in principle, also be captured by the static estimates of the effects for the second and third participation year (Table 5), we estimate the dynamic version of the effect for all lengths of participation, i.e., $\theta_2^{PP;NN}$, $\theta_3^{PPP;NNN}$, and so on. In specific, we run separate probit estimations for the different years of the program where all past intermediate outcomes (observable at each node of a sequence) and all past propensity scores are included. For instance, to calculate $\theta_2^{PP;NN}$, we estimate the probability $\Pr(PP)$ for participating in both years by estimating $\Pr(P)$ and $\Pr(PP|P)$ taking into account that $\Pr(PP) = \Pr(P) \times \Pr(PP|P)$. Analogously, the weight for individuals of sequence NN is $\Pr(P) / \Pr(NN)$.[22] Consequently, for $\theta_5^{PPPPP;NNNNN}$, we estimate five participation likelihoods per sequence.[23]

Apart from serving as a sensitivity check with respect to accounting for non-random selection patterns, the comparison with the respective dynamic treatment effects reveals their general importance. This is highly relevant because the literature has not yet sufficiently established to what extent static treatment effect estimates may be biased in the presence of voluntary changes in the treatment states of individuals over time.

Results for the dynamic treatment effects are displayed in Table 6. For the first year, the dynamic and static models are identical. For the subsequent years, just like in the static model, all estimated effects on prevention courses have a positive sign. This suggests that, in fact, in all five years, the program increases individual prevention effort. However, the estimated effect for the fourth participation year is no longer significant which points at windfall gains of participants. Moreover, the dynamic effect for the fifth year of participation is considerably larger than the corresponding static effect.

Regarding the program effects on expenses for health care, the dynamic causal model yields qualitatively similar results for the first, second, and fourth year of participation. For the third year, however, the effect – though also insignificant – is now positive. For the fifth year, we observe, in absolute terms, a smaller estimated effect which is no longer significant (p-value of

[22] Note that $\Pr(P)$ in the numerator is due to the fact that we estimate the average treatment effect on first year program participants. See Footnote 17.

[23] Since hospital information is also underreported in 2004, the number of hospital visits, the length of stay in the hospital, the total spending on hospitals, and any three-digit ICD-10 hospital discharge diagnosis are excluded from the estimation of the respective participation likelihoods (viz. Footnote 13). The first year in which these variables enters the propensity score estimation is 2005.

19 percent). This is an important difference to the static model since it points at no effect of participating in the program during five consecutive years. However, two points are important to consider. First, though pointing at a small downward bias of the static model, the differences between the static and dynamic effect is not statistically significant. Second, the insignificance of the dynamically estimated effect may simply be due to the small number of treated individuals. To exemplify this argument, we 'artifically' increase the number of observations considered in the dynamic estimation of the effect on total health care expenditures of *four* years of participation by including insurants of sequence $NPPPP$ (and $NNNNN$ as their comparison group). The intuition is that, although one year later than insurants of sequence $PPPP$, these participants are in their fourth year of participation, too. Controlling for year fixed effects, we find a negative effect of €403, an increase in the number of treated of 135 (41 percent), and, most importantly, a lower p-value of 13 percent.

— Table 6 —

4.3 Cost-benefit analysis

Since the financing of health care systems is limited, health investments should be made in the most effective way. For this reason, it is necessary to assess the cost-benefit ratio of the bonus program, making it comparable to alternative investments like, e.g., telemedicine. In order to do so, we contrast program benefits with program costs. Program benefits are the estimated reductions in yearly average total health care expenditures.[24] Program costs comprise bonus payments (80 percent), expenditures on prevention courses, personnel effort, and development costs. Unfortunately, the information about the costs for the interventions covered by the insurance company and of health care expenditures for outpatient services are unavailable (see Section 2).

[24]For the benefits of participating for five years in the program, for instance, we estimate the reduction in mean total health expenditures between 2004 and 2008.

In the upper panel of Table 7, the benefits and costs of the program for each year and participant estimated based on the static causal model are displayed. We observe positive yearly average savings in each year of participation which are significant only in the first and fifth year. In these years, for instance, they average about €185 and €310 per year, whereas program costs amount to €50 to €60 per year. The cost increase in 2006 is due to the rise in the bonus amount from €40 to €50. In the end, yearly net savings are positive for all five years of participation, yielding a benefit to cost ratio larger than one. The yearly net savings amount to €245 for those participants who participated in all five years of the program, corresponding to 5 percent of the average annual contribution of publicly insured Germans (KBV, 2009).

Testing the sensitivity of the results with respect to accounting for dynamic selection into the treatment yields substantially lower yearly net saving for participants who participated in all five years of the program (lower panel of the table). The estimated net benefit of the program reduces to about €200 and turns statistically insignificant. We observe a similar pattern in the third year. Stock et al. (2010), who focus on the effect of three years of program participation relying on the static evaluation approach might therefore report slightly upward-biased estimates. They find program net savings of €100 per year and participant. In comparison, our dynamic estimate in 2006 is roughly €50 lower and statistically insignificant, whereas our estimate of the static effect is quite close to their result.

— Table 7 —

5 Conclusion

This paper contributes to the search for policy measures that stabilize health care expenditures. We investigate a bonus program of a German social health insurer implemented in 2004 over the time horizon of five years. The bonus program aims at reducing the insurer's health care expenditures by rewarding a cash bonus for health-conscious behavior. We test whether participating in the bonus program raises preventive health behavior and reduces total health care expenditures. We also determine whether the costs associated with the bonus program are justified by its benefits.

We find that the program is effective in both increasing individual prevention effort and reducing health care spending in all years of participation. The average yearly savings are significant only in the first and fifth year. After subtracting program costs which do not include program interventions covered by the insurance company, the health insurer saves on average €245 per year for insurants who participate in all five years of the program.

Applying a dynamic causal model in order to test the sensitivity of our results with respect to accounting for dynamic selection effects, we find that the effects on net health care spending tend to be overestimated for later program years. This finding may be explained by participants no longer being included as treated observations in the analysis who exit the program after participating in obligatory screening measures on medical grounds, assuming that from then on they incur high health expenditures. Importantly, the dynamic causal model yields insignificant estimated yearly net savings in the fifth year. This could be due to the small number of the treated which makes it difficult to find a significant effect. Consequently, the finding of net savings for insurants who participate in all five years has to be regarded with caution.

Stock et al. (2010), following a static evaluation approach, find net program savings of €100 per year and insured who participate for at least three years. Our respective estimate lies in a similar range, whereas our dynamic effect is only half as large.

We conclude that health prevention represents an effective means to reduce health expenditures. With such bonus programs, health insurance companies have an effective instrument to reduce their health care expenditures in the short term. However, it is unproven whether regular health prevention over many years is equally successful. Our results, though statistically not significant, point towards higher savings for regular participants. The old adage still rings true: "An ounce of prevention is worth a pound of cure."

References

Abadie, A. and Imbens, G. W. (2006). Large sample properties of matching estimators for average treatment effects, *Econometrica* **74**(1): 235–267.

Achat, H., Mcintyre, P. and Burgess, M. (1999). Health care incentives in immunization, *Australian and New Zealand Journal of Public Health* **23**(3): 285–288.

Augurzky, B., Bauer, T. K., Reichert, A. R., Tauchmann, H. and Schmidt, C. M. (2012). Does Money Burn Fat? Evidence from a Randomized Experiment, *Ruhr Economic Papers* ***368*** .

Augurzky, B. and Schmidt, C. M. (2001). The propensity score: A means to an end, *IZA Discussion Papers* ***271*** .

Byrne, M. M. and Thompson, P. (2001). Screening and preventable illness, *Journal of Health Economics* **20**(6): 1077–1088.

Callahan, L., Mielenz, T., Freburger, J., Schreffler, J., Hottman, J., Brady, T., Buysse, K. and Schwartz, T. (2008). A randomized controlled trial of the people with arthritis can exercise program: symptoms, function, physical activity, and psychosocial outcomes., *Arthritis & Rheumatism (Arthritis Care & Research)* **59**(1): 92–101.

Charness, G. and Gneezy, U. (2009). Incentives to exercise, *Econometrica* **77**(3): 909–931.

Chatterjee, R., Venkat Narayan, K., Lipscomb, J. and Phillips, L. S. (2010). Screening adults for pre-diabetes and diabetes may be cost-saving, *Diabetes Care* **33**(7): 1484–1490.

Engelgau, M. M., Narayan, K. M. V., Thompson, T. J., Boyle, J. P., , Williamson, D. F., Flandersc, W. D., Manninen, D. L., Dong, F. B., Orians, C. E., Dasbach, E. J., Teutsch, S. M., Eastman, R., Herman, W. H. and Songer, T. J. (1998). The cost-effectiveness of screening for type 2 diabetes, *JAMA: The Journal of the American Medical Association* **280**(20): 1757–1763.

Friedrichs, M., Friedel, H. and Bädeker, W. (2009). Participant structure and economic benefit of prevention bonus programmes in company health insurance funds, *Gesundheitswesen* **71**(10): 623–627.

Giuffrida, A. and Gravelle, H. (1998). Paying patients to comply: an economic analysis, *Health Economics* **7**(7): 569–579.

Hirano, K. and Imbens, G. (2001). Estimation of causal effects using propensity score weighting: An application to data on right heart catheterization, *Health Services and Outcomes Research Methodology* **2**(3-4): 259–278.

IUHP (1999). *The Evidence of Health Promotion Effectiveness. Shaping Public Health in a New Europe. A report for the European Commission. Part 2: Evidence book*, Saint-Denis Cedex: International Union for Health Promotion and Education.

Joffe, M. M., Have, T. R. T., Feldman, H. I. and Kimmel, S. E. (2004). Model selection, confounder control, and marginal structural models: Review and new applications, *The American Statistician* **58**(4): 272–279.

KBV (2009). Statistics of the National Association of Statutory Health Insurance Physicians (KBV). **URL:** *http://www.kbv.de/index.html, Accessed: December 15, 2009*

Lechner, M. (2009). Sequential causal models for the evaluation of labor market programs, *Journal of Business and Economic Statistics* **27**(1): 71–83.

Lechner, M. and Miquel, R. (2010). Identification of the effects of dynamic treatments by sequential conditional independence assumptions, *Empirical Economics* **39**(1): 111–137.

Nichol, K. L., Nordin, J. D., Nelson, D. B., Mullooly, J. P. and Hak, E. (2007). Effectiveness of influenza vaccine in the community-dwelling elderly, *New England Journal of Medicine* **357**(14): 1373–1381.

Robins, J. (1986). A new approach to causal inference in mortality studies with a sustained exposure period–application to control of the healthy worker survivor effect, *Mathematical Modelling* **7**(9-12): 1393–1512.

Robins, J. M. and Rotnitzky, A. (1995). Semiparametric efficiency in multivariate regression models with missing data, *Journal of the American Statistical Association* **90**(429): 122–129.

Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects, *Biometrika* **70**(1): 41–55.

Rubin, D. B. (1979). Using multivariate matched sampling and regression adjustment to control bias in observational studies, *Journal of the American Statistical Association* **74**(366): 318–328.

Slater, J. S., Henly, G. A., Ha, C. N., Malone, M. E., Nyman, J. A., Diaz, S. and McGovern, P. G. (2005). Effect of direct mail as a population-based strategy to increase mammography use among low-income underinsured women ages 40 to 64 years, *Cancer Epidemiology, Biomarkers & Prevention* **14**(10): 2346–2352.

Stock, S., Schmidt, H., Büscher, G., Gerber, A., Drabik, A., Graf, C., Lüngen, M. and Stollenwerk, B. (2010). Financial incentives in the german statutory health insurance: New findings, new questions, *Health Policy* **96**(1): 51–56.

Szucs, T., Behrens, M. and Volmer, T. (2001). Volkswirtschaftliche Kosten der Influenza 1996, *Medizinische Klinik - Intensivmedizin und Notfallmedizin* **96**(2): 63–70.

Weinbrenner, S., Wörz, M. and Busse, R. (2007). *Gesundheitsförderung im europäischen Vergleich*, KomPart Verl.-Ges.

WHO (2002). *World Health Report 2002: Reducing risks, promoting healthy life*, World Health Organization.

Wu, S. (2003). Sickness and preventive medical behavior, *Journal of Health Economics* **22**(4): 675–689.

Table 1: PROGRAM DESIGN

Obligatory measures	**Remark**
Medical screening	Biennially, above the age of 35
Cancer screening	Women above the age of 20 and male above the age of 45
Check-up for children	Co-insured children below the age of 7
Elective measures (at least 3 interventions required)	**Remark**
Full vaccine protection	
Non-smoking	At least 6 months
BMI between 21 and 27	Introduced in 2007
Health promotion activities at the workplace or sickness fund	
Quality-assured services of primary prevention	
Regular exercise in sports club or fitness center	
Updated sports badge	

Notes: Unless otherwise stated, interventions are required annually.

Table 2: NUMBER OF PROGRAM PARTICIPANTS PER YEAR

	2004	2005	2006	2007	2008
Total number of participants	1,011	1,404	2,176	2,072	2,699
Share of all observations	0.017	0.024	0.037	0.037	0.046
Number of participants remaining in the program	-	603	543	403	357
Share of all observations	-	0.010	0.009	0.007	0.006

Note: Insured who exit or enter the insurer within the period 2004–2008 are not considered.

Table 3: SUMMARY STATISTICS

Variable	Mean	Std. Dev.	Min.	Max.
Expenses *(in €, 2004 – 2008)*				
Hospital *(2006 – 2008)*	486.593	2265.029	0	245946.375
Pharmaceutics	224.183	1970.748	0	236619.219
Rehabilitation	20.567	181.066	0	21446.434
Medical utilities	64.693	305.982	0	26456.537
Prevention courses	1.884	9.600	0	253.017
Inability to work payments *(in €)*	96.555	452.389	0	10819.356
East	0.143	0.350	0	1
Male	0.512	0.500	0	1
Age in years	38.160	16.060	1	94
Number of children	0.114	0.410	0	7
Having a partner	0.043	0.204	0	1
Number of other relatives	0.002	0.051	0	2
Unemployed	0.035	0.183	0	1
Pensioner	0.081	0.273	0	1
Co-insured	0.145	0.352	0	1
Type of sector				
Manufacturing	0.099	0.298	0	1
Metallic	0.073	0.261	0	1
Engineering	0.032	0.177	0	1
Social services	0.088	0.283	0	1
Administration services	0.163	0.370	0	1
Remaining services	0.172	0.378	0	1
No sectoral information	0.102	0.302	0	1
Level of education				
No graduation	0.077	0.267	0	1
Training	0.398	0.489	0	1
Senior high school	0.016	0.126	0	1
Training and senior high school	0.051	0.219	0	1
Technical college degree	0.027	0.162	0	1
University degree	0.045	0.206	0	1
No educational information	0.126	0.332	0	1
Type of occupation				
Unskilled blue-collar	0.093	0.291	0	1
Skilled blue-collar	0.107	0.309	0	1
white-collar workers	0.306	0.461	0	1
Trainee	0.097	0.296	0	1
Mini job	0.015	0.123	0	1
Midi job	0.073	0.260	0	1
No educational information, no sectoral information	0.047	0.211	0	1
Number of observations	58,618			

Note: The standard deviations over time for the expenses on hospital, pharmaceutics, rehabilitation, medical utilities, prevention courses, and the inability to work payments are 18.724, 22.032, 3.723, 5.808, 0.465, and 7.889. The unemployed, pensioners, and co-insured are excluded from the categories "no sectoral information", "no educational information" and "neither educational nor sectoral information". The shares of "unemployed", "pensioner", "co-insured", "neither educational nor sectoral information" and either the type of sector, the level of education, or the type of occupation add up to 1.

Table 4: Outcome and Covariate Comparison of Participants and Non-Participants Before the Launch of the Program

	Before Inverse Probability Weighting				After Inverse Probability Weighting				Propensity Score Probit (P vs. N)	
	Treated	Untreated	△ (Treated−Un-treated)	p-value	Treated	Untreated	△ (Treated−Un-treated)	p-value	Coeffs.	Std. Errors
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Expenses *(in €)*										
Pharmaceutics	243.114	159.993	83.121	0.009	243.114	247.850	−4.736	0.925	-	-
Rehabilitation	10.497	7.974	2.523	0.747	10.497	9.936	0.561	0.934	-	-
Medical utilities	1100.580	45.461	65.119	0.001	1100.58	1 189.399	−88.819	0,490	$10^{-7}\times 223$ **	(0.000)
Prevention courses	2.756	0.331	2.425	0.000	2.756	3.082	−0.326	0.504	0.010 **	(0.001)
Inability to work payments	149.898	115.835	34.063	0.350	149.898	122.978	26.920	0.478	-	-
Number of work inability cases *(in €)*	0.848	0.772	0.076	0.063	0.848	0.747	0.101	0.030	-	-
Number of prescribed pharmaceutics	8.215	6.385	1.830	0.000	8.215	8.295	−0.080	0.867	0.001	(0.001)
Work inability days	11.440	9.961	1.479	0.246	11.440	10.093	1.347	0.314		
East	0.178	0.143	0.035	0.002	0.178	0.172	0.006	0.667	0.104 **	(0.037)
Male	0.364	0.512	−0.148	0.000	0.364	0.362	0.002	0.906	−0.198 **	(0.031)
Age	45.684	38.061	7.623	0.000	45.684	45.797	−0.112	0.942	0.010 **	(0.001)
Number of children	0.101	0.113	−0.012	0.351	0.101	0.105	−0.004	0.778	-	-
Having a partner	0.037	0.043	−0.006	0.318	0.037	0.032	0.005	0.421	-	-
Number of other relatives	0.004	0.002	0.002	0.337	0.004	0.003	0.001	0.739	-	-
Unemployed	0.020	0.035	-0.015	0.009	0.020	0.020	0.000	0.971	−0.283 **	(0.097)
Co-insured	0.037	0.148	−0.111	0.000	0.037	0.037	0.000	0.992	−0.380 **	(0.065)
Pensioner	0.176	0.080	0.096	0.000	0.176	0.177	−0.001	0.964	0.012	(0.078)
Educational information missing	0.351	0.389	−0.038	0.015	0.351	0.353	−0.001	0.941	−0.041	(0.046)
No graduation	0.029	0.074	−0.045	0.000	0.029	0.027	0.002	0.753	−0.298 **	(0.080)
Training	0.386	0.400	−0.014	0.380	0.386	0.388	−0.002	0.938	-	-
Senior high school	0.017	0.016	0.001	0.877	0.017	0.014	0.002	0.530	-	-
Training and senior high school	0.080	0.050	0.030	0.000	0.080	0.083	−0.003	0.768	0.188 **	(0.054)
Technical college degree	0.043	0.027	0.016	0.002	0.043	0.043	0.000	0.955	0.155 **	(0.071)
University degree	0.094	0.044	0.050	0.000	0.094	0.092	0.002	0.843	0.280 **	(0.053)
Sectoral information missing	0.364	0.364	0.000	0.996	0.364	0.362	0.002	0.904	-	-
Manufacturing	0.037	0.098	−0.061	0.000	0.037	0.038	−0.001	0.861	0.024	(0.107)
Metallic	0.036	0.074	−0.038	0.000	0.036	0.035	0.001	0.897	0.108	(0.109)
Engineering	0.048	0.032	0.015	0.006	0.048	0.048	0.000	0.977	0.200 *	(0.103)
Remaining services	0.146	0.172	−0.026	0.027	0.146	0.146	−0.001	0.965	−0.006	(0.086)
Social services	0.124	0.087	0.036	0.000	0.124	0.124	0.000	0.970	0.060 **	(0.090)
Administration services	0.240	0.163	0.077	0.000	0.240	0.238	0.002	0.912	0.124	(0.086)
Other	0.007	0.010	−0.003	0.353	0.007	0.01	−0.003	0.406	-	-
Unskilled blue-collar workers	0.021	0.091	−0.070	0.000	0.021	0.022	−0.001	0.870	−0.519 **	(0.108)
Skilled blue-collar workers	0.047	0.108	-0.062	0.000	0.047	0.046	0.000	0.949	-0.367 **	(0.100)
white-collar workers	0.380	0.306	0.074	0.000	0.380	0.381	−0.001	0.954	−0.119	(0.077)
Mini job	0.034	0.015	0.019	0.000	0.034	0.032	0.002	0.782	0.152	(0.102)
Midi job	0.128	0.072	0.055	0.000	0.128	0.128	−0.001	0.953	−0.028	(0.084)
Trainee	0.095	0.098	−0.003	0.786	0.095	0.106	−0.010	0.308	-	-
Educational and occupational information missing	0.296	0.310	−0.014	0.357	0.296	0.285	0.011	0.522	-	-
Constant									−2.406 **	(0.076)
Number of observations	1,011	57,607			1,011	57,607			58,618	

Notes: Descriptive statistics are computed on the common support. p-values acknowledge the randomness of the weights, see Section 3. For the controls, statistics for reference categories explicitly reported. Probit Estimation: ** significant at 5%; * significant at 10%. ICD-10 work inability diagnoses that passed selection criterion for probit estimation and were included as control dummy variables (+ (-) denotes that the respective variable has a positive (negative) coefficient in the index of the participation probability): + relapsing fevers (A68) *, + dystonia (G24)**, + nonsuppurative otitis media (H65)**, + hypertensive heart disease with (congestive) heart failure (I11)**, + bronchitis not specified as acute or chronic (J40)**, + allergic contact dermatitis (L23)**, + sunburn (L55)**, + labour and delivery complicated by fetal stress (O68)**, + other maternal diseases classifiable elsewhere but complicating pregnancy, childbirth and the puerperium (O99)**, + other congenital malformations of limbs (Q74)**, + haemorrhage elsewhere not classified (R58)**, + fracture of shoulder and upper arm (S42)**, + superficial injuries involving multiple body regions (T00)**.

Table 5: (Static) Estimation of the Program Effect on Expenses for Prevention Courses and Total Health Care

	$\theta_1^{P;N}$	$\theta_2^{PP;NN}$	$\theta_3^{PPP;NNN}$	$\theta_4^{PPPP;NNNN}$	$\theta_5^{PPPPP;NNNNN}$
	2004	2005	2006	2007	2008
Prevention Courses					
Untrimmed	3.936 **	1.762 **	2.972 **	5.750 **	5.843 **
	0.000	*0.000*	*0.001*	*0.000*	*0.000*
Trimmed ‡	3.890 **	1.768 **	2.316 **	5.712 **	5.829 **
	0.000	*0.000*	*0.001*	*0.000*	*0.000*
Total Health Care†					
Untrimmed	-184.586 **	-108.374	-258.802	-307.513	-524.107 *
	0.033	*0.417*	*0.157*	*0.284*	*0.050*
Trimmed‡	-187.271 **	-109.194	-264.720	-302.600	-525.422 **
	0.027	*0.414*	*0.153*	*0.295*	*0.050*
Number of Treated	1 010	603	543	403	357
Non-Treated	57 367	56 806	54 791	55 224	54 489

Notes: Comparison of the covariate adjusted weighted mean outcomes between participants and non-participants. Covariate selection rule for regression adjustment as well as propensity score estimation applied. p-values (in italics) acknowledge the randomness of the weights (see Section 3). † Expenses for prevention courses excluded. ‡ Observations with the largest five-percent weights removed. ** Significant at 5%; * significant at 10%.

Table 6: Dynamic Estimation of the Program Effect on Expenses for Prevention Courses and Total Health Care

	$\theta_1^{P;N}$	$\theta_2^{PP;NN}$	$\theta_3^{PPP;NNN}$	$\theta_4^{PPPP;NNNN}$	$\theta_5^{PPPPP;NNNNN}$
	2004	2005	2006	2007	2008
Prevention Courses					
Untrimmed	3.936 **	2.493 **	2.750 *	1.443	7.050 **
	0.000	*0.001*	*0.065*	*0.328*	*0.000*
Trimmed‡	3.890 **	1.694 **	5.083 **	0.735	6.685 **
	0.000	*0.008*	*0.000*	*0.595*	*0.000*
Total Health Care†					
Untrimmed	-184.586 **	-196.409	31.418	-307.650	-469.391
	0.033	*0.380*	*0.913*	*0.385*	*0.193*
Trimmed‡	-187.271 **	-161.016	-124.492	-347.881	-440.782
	0.027	*0.476*	*0.674*	*0.345*	*0.249*
Number of Treated	1 010	597	502	329	253
Non-Treated	57 367	53,541	52,476	50,875	49,546

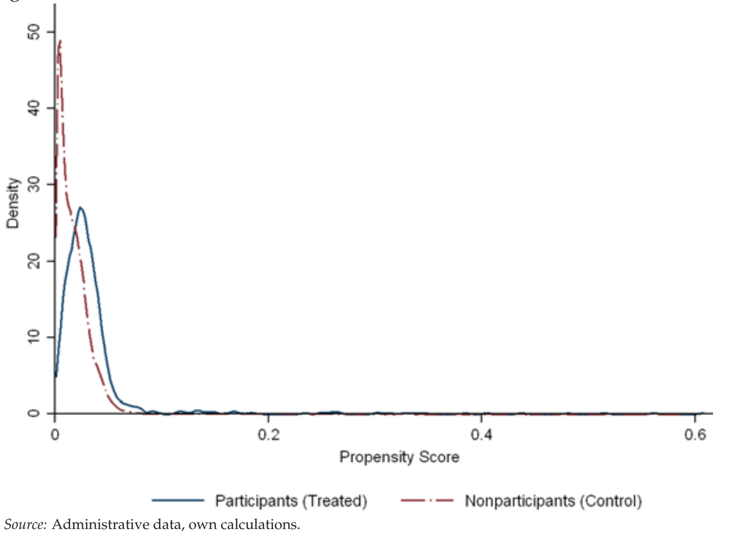
Notes: Comparison of the covariate adjusted weighted mean outcomes between participants and non-participants. Covariate selection rule for regression adjustment as well as propensity score estimation applied. p-values (in italics) acknowledge the randomness of the weights (see Section 3). † Expenses for prevention courses excluded. ‡ Observations with the largest five-percent weights removed. ** Significant at 5%; * significant at 10%.

Table 7: COST-BENEFIT ANALYSIS

	$\theta_1^{P;N}$	$\theta_2^{PP;NN}$	$\theta_3^{PPP;NNN}$	$\theta_4^{PPPP;NNNN}$	$\theta_5^{PPPPP;NNNNN}$
	2004	2005	2006	2007	2008
STATIC MODEL					
Savings	184.59 **	84.41	150.15	214.29	310.02 **
	0.003	*0.384*	*0.156*	*0.118*	*0.046*
Costs	52.98	50.81	62.02	64.80	64.89
Net Savings	131.61	33.60	88.13	149.49	245.13
Number of Treated	1,010	603	543	403	357
Non-Treated	57,367	56,806	54,791	55,224	54,489
DYNAMIC MODEL					
Savings	184.59 **	208.96	111.66	141.55	199.53
	0.003	*0.188*	*0.505*	*0.483*	*0.313*
Costs	52.98	51.54	61.80	60.49	66.55
Net Savings	131.61	157.42	49.86	81.06	132.98
Number of Treated	1,010	597	502	329	253
Non-Treated	57,367	53,541	52,476	50,875	49,546

Source: Administrative data, own calculations.
Note: Savings and costs reflect yearly averages. Average savings are estimated employing the DiD-IPW estimator with regression adjustment. Development costs (€ 15,000) incurred in equal parts in 2003 and 2004. With exception of the bonus payments (remaining) costs in the amount of € 250,982 are divided in equal shares among the five years. Information on fifth year program costs not available. We assume that the costs are the same as in the previous year. The increase in the expenses on prevention courses are included into the costs. *p*-values (in italics) acknowledge the randomness of the weights (see Section 3).** Significant at 5%; * significant at 10%.

Figure 1: PROPENSITY SCORE DENSITY FOR PROGRAM PARTICIPATION IN 2004



Source: Administrative data, own calculations.
Note: Density estimates of the probability to participate in the program base on all individuals that were insured with the health insurer at least 350 days each year between 2004 and 2008.

Appendix

Stepwise Estimation Procedure

A Computation of Inverse-Probability Weights

1. Conduct bivariate regressions of program participation indicator P on $X_k \in X \; \forall \, k$, where $k = 1, ..., K$ is the total number of explanatory variables.

2. Select X_k whose estimated coefficients have a p-value less or equal to 0.05. Denote this covariate set as X^*.

3. Conduct a multivariate regression of P on X^*: $P_{i,t} = f(X^*_{i,t-1}) + e_{i,t}$, where $f(\cdot)$ is a linear-in-parameter function.

4. Compute the predicted probability of participation, $\hat{e}(X^*_{i,t-1})$.

5. Compute the weight: $w_{i,t} = P_{i,t} + (1 - P_{i,t}) \times \frac{\hat{e}(X^*_{i,t-1})}{1-\hat{e}(X^*_{i,t-1})}$.

B Combining Weighting with Regression Adjustment

1. Conduct a trivariate regression of the first difference of the outcome variable ($\Delta Y_{i,t} \equiv Y_{i,t} - Y_{i,t-1}$) on $w_{i,t} \, [\beta g(X_k) + \gamma P_{i,t}] \; \forall \, k$.

2. Select X_k whose estimated coefficients have a p-value less or equal to 0.05. Denote this covariate set as $\tilde{X}$.

3. Regress $\Delta Y_{i,t}$ on $w_{i,t} \left[\beta g(\tilde{X}_{i,t-1}) + \gamma P_{i,t}\right]$.

C The estimate of the intention-to-treat effect is the estimated coefficient $\hat{\gamma}$.