RUHR

ECONOMIC PAPERS

Alfredo R. Paloyo

Co-pay and Feel Okay: Evidence of Illusory Health Gains from a Health Insurance Reform

Ruhr Graduate School in Economics

#142

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 #142

Responsible Editor: Thomas K. Bauer

ISSN 1864-4872 (online) – ISBN 978-3-86788-160-9

Ruhr Economic Papers #142

Alfredo R. Paloyo

Co-pay and Feel Okay: Evidence of Illusory Health Gains from a Health Insurance Reform

Bibliografische Informationen der Deutschen Nationalbibliothek

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

ISSN 1864-4872 (online)
ISBN 978-3-86788-160-9

Alfredo R. Paloyo[1]

Co-pay and Feel Okay: Evidence of Illusory Health Gains from a Health Insurance Reform

Abstract

The reliability of general self-rated health status is examined using the reform of the public health insurance system of Germany in 2004 as a source of exogenous variation. Among others, the reform introduced a co-payment for ambulatory doctor visits and increased the co-payments for prescription drugs. This natural experiment allows identification of the causal impact of the program on self-assessed health and hence reveals the sensitivity of this subjective measure to a perturbation in the insurance system. Using data from the German Socio-Economic Panel, the results indicate that after the policy intervention, the respondents in the treated group perceived their own health status as better than their hypothetical untreated state even when there is no discernible impact on actual health.

JEL Classification: G22, H43, I18

Keywords: natural experiment, cognitive dissonance, self-rated health status

October 2009

1 RGS Econ, RUB, and RWI. – For their valuable contributions to the improvement of this paper, I thank the following: Thomas K. Bauer, Ronald Bachmann, Daniel Baumgarten, Christoph Ehlert, Simon Golshan, Tobias Körner, and Hendrik Schmitz. Participants of the RGS Workshop at RWI in Essen, the 23rd Annual Conference of the European Society for Population Economics in Sevilla, the 24th Annual Congress of the European Economic Association in Barcelona, and the 2009 Annual Conference of the European Association of Labour Economists in Tallinn are gratefully acknowledged. All remaining shortcomings are my responsibility. The Leibniz Association through the RGS Econ provided financial support. *– All correspondence to Alfredo R. Paloyo, Ruhr Graduate School in Economics, Hohenzollernstr. 1–3, 45128 Essen, Germany, e-mail: alfredo.paloyo@rwi-essen.de.*

1 Introduction

Measures of subjective well-being are now extensively used in many areas of empirical economics. These measures are particularly prevalent in studies that deal with general life satisfaction [Winkelmann and Winkelmann 1998; Kassenböhmer and Haisken-DeNew 2009], job satisfaction [Hamermesh 2001], and health [Bago d'Uva et al. 2008; Bago d'Uva, O'Donnell and van Doorslaer 2008]. At least two reasons give rise to the wide availability of this type of data in surveys. One is that it is relatively inexpensive to collect and the other is that it is fairly simple for both the interviewer and the respondent to understand. Questions eliciting subjective assessments are typically short and simple such that the respondents are forthcoming with their answers. For example, current health status is self-assessed using a five-point scale.[1] This is a rather cheap alternative to extracting a blood sample and having it tested in a qualified medical laboratory, a point also made by Hamermesh [2004]. In the latter, the interviewer should be a trained medical technologist or at least accompanied by one. There is also the possibility that the respondent may not cooperate at all considering the level of intrusion a blood extraction entails. The attractiveness of subjective measures to survey designers is therefore understandable.

Despite some skepticism [Bertrand and Mullainathan 2001], these measures have—at least recently—proved to be enticing to economists and other social scientists as well. The use of these variables in social studies has become so popular primarily because the wide availability of the data lends itself to analysis, if not invites it. Hamermesh [2004] calls this the "Mt. Everest phenomenon".[2] Happiness or life-satisfaction research exemplified by Easterlin [1995], for example, has enjoyed a boom in economics.[3] Economists now routinely analyze the impact of certain exogenous changes

[1] The typical question is "How would you describe your current health?". A respondent can choose among the following answers: "very good", "good", "satisfactory", "poor", and "bad".

[2] Asked why he wanted to climb Mt. Everest, George Mallory, an English mountaineer who died scaling the mountain, famously retorted, "Because it's there."

[3] For a survey, see Frey and Stutzer [2002].

on one or another subjective outcome, e.g., the effect of unemployment on either life satisfaction [Kassenböhmer and Haisken-DeNew 2009] or health, measured as self-assessed health [Böckerman and Ilmakunnas 2009].

Self-rated health status (SRHS) is perhaps one of the most common subjective measure that one finds in the literature. This is used both in econometric as well as in epidemiological research. Since the study of Mossey and Shapiro [1982], SRHS has been generally accepted as a good predictor of mortality and morbidity, particularly in the epidemiological literature. However, the reliability of SRHS is obviously not unquestionable. Using Australian data, Crossley and Kennedy [2002] show that SRHS is sensitive to the type of questions that preceded it—what Bertrand and Mullainathan [2001] call the "ordering effect". As an example, if previous questions elicited good memories, then the respondent is more likely to rate her health as good. Bound [1991] points out the endogeneity problem that arises out of using SRHS in labor-force-participation studies. Retirement decisions made for reasons other than health (say, a person could not get along with co-workers), for example, may be rationalized by people by instead claiming that they were no longer able to work due to health reasons, which carries with it less stigma than some other possible reasons. On top of the problem induced by measurement error, self-assessed health is also subject to reporting heterogeneity. For example, Bago d'Uva et al. [2008] show that subjective measures suffer from differential reporting on the basis of the level of education and income. Thus, this measure of health is apparently not very stable.

This paper contributes to the literature on the reliability of SRHS specifically and measures of subjective well-being in general by exploiting a natural experiment in the health insurance system. In 2004, a major health insurance reform was adopted in Germany that effectively raised the price of medical services. We estimate the causal impact of this reform on SRHS. The reason we expect the reform to have an effect on SRHS—irrespective of whether or not there was a concomitant appreciation of objective health—is related to at least two strands of the literature. The first involves the problem of "cognitive dissonance"; the other is related to the informational effect of prices. These are explained in more detail in Section 2.

In this paper, we find that the reform had a positive and significant (albeit modest) effect on SRHS, i.e., that the reform made people perceive themselves as healthier. However, this perception of one's health improvement is not accompanied by a corresponding improvement of actual or objective health. Thus, we show that a subjective measure of health such as SRHS fluctuates even without any change in objective measures of health. That is, there are other factors that influence SRHS that are presumably unrelated to a person's actual health status.[4] While SRHS is traditionally thought of as a function of objective health plus some other factors, what we show here is that one of these factors includes the price of health services through either its signaling effect or cognitive dissonance and not merely through its effect on the quantity demanded of health services.

We therefore provide another piece of evidence that invites caution in the use of these subjective measures, particularly of SRHS. The implication is that survey designers should collect more objective measures of health if this is cost-efficient. In this regard, the use of biomarkers (biological indicators) in data-gathering have become increasingly common even in developing countries.[5] Moreover, health researchers would be better off using such measures if they are available. While we are far from saying that subjective measures are meaningless[6], we do emphasize that research using these measures should always be taken with a grain of salt.

[4]Whether or not this improvement in self-assessed health is important in and of itself is an issue left for further research. For example, does the fact that one thinks of oneself as healthy positively contribute to one's actual health?

[5]Biomarkers include measures obtained from blood samples as well as simple anthropometrics, such as height, weight, and waist circumference. Other routine biomakers that do not involve extremely intrusive methods include the systolic and diastolic blood pressures.

[6]In fact, they are likely to be pregnant with meaning such that disentangling the components of, say, general life satisfaction is considered to be a worthwhile endeavor by some. However, the extent to which economists can contribute to the literature is unclear. For a discussion, see Hamermesh [2004].

2 Theoretical considerations

Eyster [2002] develops a model where economic agents have a taste for consistency. In the model, people experience what psychologists call cognitive dissonance, a phenomenon wherein the agent changes her beliefs over the utility she derives from her actions. A related concept is "confirmatory bias", where new information is processed by agents in such a way as to confirm currently held beliefs [Yariv 2005]. Among other reasons, cognitive dissonance is triggered by the "sunk-cost effect". An example is somebody who stays until the end of a movie that she realizes to be horrible five minutes in because she has already paid for the ticket. While standard economic theory would have us believe that the sunk cost of the ticket should not factor into the decision of the viewer to stay or go, it is not unusual for people to rationalize the sunk cost by holding out until the end credits have started rolling. This is why the phenomenon is sometimes referred to as "escalation of commitment" or "entrapment" [Eyster 2002; Ashraf, Berry and Shapiro 2007]. The assumption here is that the act of paying for something that turns out to be non-enhancing to utility has itself a disutility apart from the cost price (e.g., the feeling of regret). In order to avoid this extra cost, people "rationalize the past" for their actions to be consistent.

It is possible that the health insurance reform triggered the phenomenon of cognitive dissonance. As part of the reform package, a co-payment for doctor visits was introduced which effectively raised the price of health services. While the quality of medical services may have remained the same, the fact that people are now paying for something that used to be free may induce them to think that their health is actually better than before. This allows them to avoid the psychological cost of paying for nothing.

The change in the effective price could also have an impact on SRHS due to the informational effect of prices. As early as 1944, Scitovszky already pointed out that people may infer the quality of a product or service from its price. He says, "A commodity offered at a lower price than competing commodities will be both more attractive to the consumer on account of its greater cheapness and less attractive on account of its suspected inferior quality." Evidence of this type of behavior wherein price is used as a signal for the quality of the good is observed in many markets, e.g., "fountain-pen ink and car wax and vodka, skis, and television sets" [Bagwell and Riordan 1991 and the references therein]. This is also seen in the wine market [Gergaud and Livat 2007] and there is some evidence that consumers also use price as a signal of quality in the market for genetically-modified food [Hwang, Roe and Teisl 2006].[7]

The phenomenon is more prevalent in markets where there is a high degree of information asymmetry between sellers and buyers (as in the case between doctors, who provide medical service, and patients, who demand it). To ameliorate the asymmetry, sellers may use a high price to signal to the consumer that their wares are of high quality, with the assumption that high-quality goods are more expensive to produce. If consumers observe an increase in price after the reform, they may associate this with an increase in the effectiveness of the service.[8]

Especially relevant is the study of Shiv, Carmon and Ariely [2005]. In an experiment, they show that people who pay the discounted price for a particular product (in this case, an energy drink) derive less benefit from it than people who pay the full price. They call this the "placebo effect" which works at the subconscious level. Since the reform of the insurance system in Germany effectively raised the price of health services, consumers may experience a perceptible change in its efficacy. They may therefore rate themselves as healthier since the price change is associated with an improvement in the perceived effectiveness of the service.

[7] This issue is related to Veblen goods, named after Thorstein Veblen, who noted the phenomenon of conspicuous or ostentatious consumption. For Veblen goods, a price increase is associated with an increase in quantity demanded but the mechanism by which this occurs is not through consumer optimization given a budget constraint—as in Giffen goods—but rather through an interaction between prices and preferences. [Leibenstein 1950]

[8] The conditions under which the signal is credible, i.e., where the low-quality producers cannot successfully imitate the signal of the high-quality producers, is derived in the seminal work of Akerlof [1970]. See also Bagwell and Riordan [1991]. In the text, the consumer is comparing two goods from different time periods. In standard models, a consumer has two types of goods in the same period. However, this is just a change in the frame of reference.

3 Institutional background

The German health insurance system is practically divided into two independent parts: the public or statutory component and the private component. The former covers roughly 90 percent of the population while the latter accounts for most of the rest. Within these two systems are dozens of providers, although those belonging to the statutory health insurance (SHI) system are highly regulated and hence are hardly differentiated. The option of individuals to choose from either one of these systems is legally regulated. Notably, employees whose salary falls below an income threshold (which stood at €3,975 per month in 2007) cannot opt out of the public insurance system[9]. Those who earn more can choose between being insured under the public system or the private system. The self-employed are also freely able to choose between the two systems.

While not set in stone, the placement of an individual in one of these systems is almost permanent. Incentives are in place that effectively discourage switching between the two systems and, presumably, multiple switching in one person's lifetime is rare. This strict division is exploited in this paper to identify the causal effect of a health insurance reform on SRHS. The reform package described in the next section affected only those who are in the SHI system, effectively creating a treatment group and a control group, the latter principally composed of the privately insured. The policy handle in this institutional setting is exogenous and there is no ambiguity in the direction of causation.

On 1 January 2004, the Statutory Health Insurance Modernization Act[10] (GMG) came into force in Germany.[11] The major components of the Act are the introduction

[9]In German, this is *Gesetzliche Krankenversicherung*, which is why the common translation into English is "statutory health insurance" instead of merely "public". It is not simply that one is barred from entering the private system instead of the public; it is that one is mandated or compelled to be insured under the public system. This is the reason that the fraction of uninsured people in Germany is negligible. In this paper, the terms "public insurance" and "statutory insurance" are used interchangeably.

[10]In German, *Gesetz zur Modernisierung der gesetzlichen Krankenversicherung* or simply GMG for *Gesundheitsmodernisierungsgesetz*.

[11]There was also a reform in the hospital sector of Germany in 2004. Hospitals shifted away from a

of co-payments for ambulatory doctor visits, an increase in the co-payments for prescription medicine, and the exclusion of over-the-counter (OTC) drugs from insurance coverage. This set of policy interventions applied to the publicly insured who are at least 18 years old. The explicit objective of the GMG is to reduce public health care costs by manipulating the price faced by consumers of health care services.

Due to the GMG, people who are part of the SHI system are now required to pay € 10 for their first doctor visit per calendar quarter (with a few exemptions). Subsequent visits to the same doctor or to another doctor—provided a referral from the first doctor is obtained—in the same quarter are not charged the co-payment. Emergency visits to the hospital are also subject to the same scheme. That is, despite having paid € 10 for a first-time visit to a doctor, one is also charged another € 10 if one ends up in the emergency room within the same quarter. As regards prescription drugs, co-payments now range from € 5 to € 10, depending on the size of the drug package. Some OTC drugs that were previously prescribed anyway by the doctor were paid for by the insurance company. The recent reform put a stop to this practice. Now, all OTC drugs are paid for by the consumer in full.

Potentially, total expenditure just for the co-payments for ambulatory care could be € 120 per year for a three-person household whose members are all publicly insured and are not exempt from co-payments, accounting for less than 1 percent of the mean annual household net income of the lowest income quintile of people in the SHI system (€ 15,524).[12] Moreover, the schedule of co-payments is constructed in such a way that people are able to control the total amount of co-payments by keeping all doctor visits in one quarter. It is no surprise therefore that studies looking at the effect of the increase in co-payments for ambulatory doctor visits on the number of such visits of

cost-plus reimbursement system to a prospective-payment system. In the new system, hospitals receive a fixed payment for the treatment of a patient. This introduced incentives for hospitals to contain costs [Schwierz 2009]. However, this new system is more likely to negatively contribute to a patient's health. To the extent that the mapping from objective health to self-assessed health is monotonic, we should therefore expect SRHS to decline, *ceteris paribus*, after the hospital reform. The positive and significant effect we present in Section 5.2 may be thus said to be somewhat understated.

[12]This is the mean for the lowest income quintile of those who are publicly insured in the constructed sample used in this paper. The details of how the sample was selected are described in Section 4.

the publicly insured conclude that it failed to achieve its stated objective [Augurzky, Bauer and Schaffner 2006; Schreyögg and Grabka 2008].

4 Data construction and description

The data used for the present analysis come from the German Socio-Economic Panel (SOEP).[13] The SOEP is a representative and longitudinal annual survey that started in 1984. In 2008, more than 20,000 persons were sampled. Further details can be found in Wagner, Frick and Schupp [2007].

The waves 2000–2003 and 2005–2007 (seven waves) of the SOEP are used here (i.e., four years before and three years after the reform became effective). The data collected in 2004 were discarded because of the way the questions on the number of doctor and hospital visits were asked. The questions are, "Have you gone to a doctor within the last three months?" and "Were you ever admitted to a hospital for at least one night in the previous year?" Since most interviews are conducted in the first quarter of the year, the answers from 2004 straddle both the pre-intervention and post-intervention periods. To be consistent with the other papers that look at the GMG, those who are older than 65 years are also dropped. Civil servants were also dropped from the sample because of their special insurance status.[14]

We also dropped individuals who changed from public to private insurance and vice versa (7,251 observations). The intention of deleting these observations is to keep the sample unadulterated: everyone in the sample remains in the same group before and after the policy intervention. The concerns over substitution and dropout biases [Heckman et al. 2000], wherein either members of the control group seek the treatment

[13] The data used in this paper were extracted using the Add-On package PanelWhiz v2.0 (Nov 2007) for Stata®. PanelWhiz was written by Dr. John P. Haisken-DeNew <john@panelwhiz.eu>. The PanelWhiz generated DO file to retrieve the SOEP data used here and any Panelwhiz Plugins are available upon request. Any data or computational errors in this paper are my own. Haisken-DeNew and Hahn [2006] describe PanelWhiz in detail.

[14] Civil servants (*Beamte* in German) receive state assistance which roughly covers 50 percent of their expenditures for health care services.

on their own or members of the treatment group drop out of treatment, are thus mitigated. Ultimately, the operational sample contains 89,897 person–year observations.[15]

When randomization is successfully carried out, observational units in the treatment and control groups can be expected to be similar at least across observable characteristics at baseline (i.e., pre-intervention). The goal is to have a control group that is a suitable counterfactual to the treatment group with the only difference being the random assignment into treatment. It is therefore prudent to examine the differences in means between the statutorily insured and the privately insured in terms of important characteristics. This is presented in Table 1, where it is shown that the publicly insured are different from the privately insured. We therefore control for these observable differences by including these variables in the regression analyses.

5 Identification, estimation, and discussion

5.1 Identification strategy

Consider the linear DID estimation strategy, which can be represented as a two-way fixed-effects regression model:

$$SRHS_{it} = \alpha + \delta S_i + \tau \left(S_i \times P_t\right) + \boldsymbol{\beta}'\mathbf{X}_{it} + \boldsymbol{\theta}'\mathbf{Z}_t + u_{it}, \tag{1}$$

where $SRHS_{it}$ is the SRHS of individual i in year t, S_i is a treatment-group indicator that takes on the value of 1 if individual i is statutorily insured, P_t is a dummy variable that is equal to 1 if the observation is from 2005 or later, $\mathbf{X}_{it}$ is a vector of control variables, $\mathbf{Z}_t$ is a vector of person-invariant year fixed effects, and u_{it} is a stochastic

[15]Observations that do not contain information in any of the covariates used in the regressions are also dropped with the assumption that these pieces of information are missing completely at random. Retaining the insurance system switchers did not substantially affect the results described below. Moreover, with their inclusion, we are able to estimate the model with individual-level fixed effects (FE). The results presented in Section 5.2 hold as well in the conditional FE ordered logit model [Chamberlain 1980; Ferrer-i-Carbonell and Frijters 2004].

TABLE 1 — EQUALITY-OF-MEANS TEST, PRE-INTERVENTION OBSERVATIONS

Variable	Privately insured	Publicly insured	Difference	Standard error
SRHS	3.720	3.541	0.179	0.015
Female	0.342	0.528	−0.186	0.008
Age	44.679	40.596	4.082	0.201
Years of education	14.259	11.955	2.304	0.039
Living with children under 16	0.392	0.417	−0.026	0.008
Number of children	0.679	0.736	−0.057	0.016
Living in West Germany	0.863	0.747	0.116	0.007
Immigrant	0.063	0.152	−0.090	0.006
(log) Household net income	8.521	8.079	0.443	0.009
(log) Household labor income	11.049	10.400	0.650	0.014
White-collar worker	0.363	0.393	−0.030	0.008
Blue-collar worker	0.021	0.274	−0.253	0.007
Married	0.695	0.648	0.048	0.008
Widowed	0.007	0.016	−0.009	0.002
Divorced	0.078	0.062	0.016	0.004
Separated	0.017	0.017	0.000	0.002
Full-time employed	0.679	0.525	0.153	0.008
Part-time employed	0.048	0.139	−0.091	0.006
Self-employed	0.386	0.038	0.348	0.004
Not currently working	0.218	0.245	−0.026	0.007
Partner died recently	0.001	0.002	−0.001	0.001
Recent separation from partner	0.018	0.017	0.001	0.002
Recently divorced	0.006	0.006	0.000	0.001
Had a child born	0.041	0.038	0.003	0.003
Disabled	0.048	0.067	−0.019	0.004
Smoker	0.262	0.311	−0.049	0.008
Degree of handicap	2.596	3.878	−1.281	0.247

SOURCE: Own computation based on SOEP 2000–2003. Numbers were rounded off to the nearest thousandths.

disturbance term. We refer to $(S_i \times P_t)$ as the DID variable and τ as the DID coefficient. The parameter of interest is τ, which is the average treatment effect of the policy on the outcome variable for the treated group (ATET for "average treatment effect on the treated").[16] Estimating τ from Equation (1) removes the bias associated with the permanent difference between the treatment and control groups.

A complication is the ordinal nature of the dependent variable: SRHS is measured in a scale from 1 to 5, where 1 is "bad" and 5 is "very good".[17] One can straightforwardly estimate a pooled ordered logit (POL) model to take this into account. The reported SRHS can be thought of as a discretized realization of an unobserved contin-

[16] Henceforth, the (average) treatment effect mentioned here refers to the ATET.
[17] The original scaling in the dataset is reversed (i.e., 1 is "very good").

uous variable that represents the perception of one's health, $SRHS^*$, which is assumed to have a linear index: $SRHS^*_{it} = \gamma'\mathbf{W}_{it} + \varepsilon_{it}$, where $\mathbf{W}_{it}$ is a vector of covariates and ε_{it} is a typical error term that is assumed to have a logistic distribution. The reported SRHS is linked to its latent counterpart by the following rule:

$$SRHS_{it} = j \text{ if } \phi_{j-1} < SRHS^*_{it} \leq \phi_j,$$

where $i = 1, \ldots, N$, $t = 1, \ldots, T_i$, and $j \in \{1, 2, 3, 4, 5\}$. The ϕ_j's are unknown cutpoints or threshold parameters that have to be simultaneously estimated with γ, which contains the parameter of interest (the causal effect of the GMG on SRHS). As usual, it is assumed that $\phi_0 = -\infty$ and $\phi_5 = \infty$. This is implemented in a DID framework by simply replacing the linear index function in the ordered logit model with the index provided in Equation (1) minus the constant α.

One can alternatively give a cardinal interpretation to the dependent variable and then estimate an ordinary least-squares (OLS) model. OLS has the advantage of providing a simple way to calculate the ATET since the estimated coefficients can directly be interpreted as the treatment effect. The obvious limitation here is that the nonlinear conditional expectation function is only linearly approximated. Nevertheless, we estimated both a linear and a nonlinear model.

5.2 Estimation results

Table 2 presents the results following the estimation of the POL and OLS models, respectively. Columns (1) and (3) present results with no socioeconomic control variables (i.e., the model is estimated only with a treatment indicator, the DID variable, quarter dummies, and year fixed effects). Columns (2) and (4) include the following control variables: sex, age (and its square), years of schooling, whether a child below 16 years old lives in the household, the number of children living in the household, (log) household net income, (log) household labor income, dummy variables for white- or blue-collar workers, dummy variables for civil status (married, widowed, divorced, or separated), dummy variables for employment status (full-time, part-time, self-employed, or not currently employed), the number of days the respondent was unable to work in the previous year, dummy variables for life shocks (partner recently died, recently separated from partner, recently divorced, recently had a child born), whether the respondent is disabled, the degree of disability, whether the respondent is a smoker, whether it was an oral interview, whether the respondent lives in West Germany, whether the respondent visited a doctor in the previous quarter, whether the respondent visited a hospital in the previous year, and whether the respondent is an immigrant.

TABLE 2 — REGRESSION RESULTS

	POL		OLS	
	(1)	(2)	(3)	(4)
DID coefficient	0.148***	0.121**	0.058***	0.042**
	[0.047]	[0.050]	[0.022]	[0.020]
Controls	None	Full	None	Full
(χ^2-) F-statistic	(252.41)	(11,276.20)	22.47	328.09
(Pseudo) R^2	(0.002)	(0.105)	0.005	0.244
Observations	89,897		89,897	

NOTES: Bracketed numbers are standard errors clustered at the individual level. Regressions include quarter dummies and year fixed effects. See text for the list of other control variables. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
SOURCE: Own computation based on SOEP 2000–2003 and 2005–2007.

In all four specifications, the estimates of the DID coefficient come up as positive and significant. The estimates are also robust to the inclusion of socioeconomic control variables although their inclusion reduces the magnitude of the coefficient estimates. The estimated DID coefficients from OLS regressions can be directly interpreted as the ATET, which ranges from 0.042 to 0.058, and these estimates are all significiant at least at the 5-percent level. The standard errors are clustered at the individual level since the observations are pooled and therefore it is unlikely that observations for the same individual in different years are independent of each other. To address this issue, we allow for an arbitrary covariance structure within each cluster (i.e., each person).

Based on the estimated treatment effects, one could say that the impact is by no means large. However, it is robust to both changes in the functional form[18] and the inclusion of additional covariates. While the magnitude of the estimated coefficient becomes smaller when using the full set of control variables, this only means that we are properly accounting for possible differential time effects that are functions of covariates as discussed in Section 5.4. To get an idea of the size of this effect, we compare the treatment effect with the estimated coefficient of (log) household net income, which is 0.092 with a standard error of 0.010. Using the treatment effect in Column (4), we find that the reform is equivalent to raising household net income by 0.45 log points or by about €1.57 per month.

Overall, the results indicate that there was definitely an improvement in the SRHS of the publicly insured in Germany after the GMG of 2004. Whether this was purely a psychological effect such as cognitive dissonance or the result of an actual improvement of health status is discussed in the next section.

5.3 Objective health measures

Studies of causal effects do not necessarily identify the mechanism through which an action generates a reaction. Isolating the pathways through which a policy affects an outcome variable of interest requires more than parameter estimates of treatment effects. For example, there are multiple ways through which the GMG can affect SRHS. It could be through an actual improvement in health or merely a psychological effect that has nothing to do with one's objective health status. We therefore examined whether there was an improvement in the actual health of the publicly insured after the reform of 2004. If we find evidence that objective health measures improved, then the positive and significant effect on SRHS of the health insurance modernization reform that we found in Section 5.2 would not be puzzling at all. In the absence of any

[18] Although, technically, the estimated coefficients from the nonlinear DID models cannot be interpreted as treatment effects. Nevertheless, Puhani [2008] shows that the sign and significance of the coefficient estimate carry over to the ATET if one undertakes its computation.

such objective improvement, the change in SRHS is likely the result of a mere change in preferences of the consumers of health services in response to the effective price increase or the result of using the increased price as a signal of higher-quality (and hence effectivity) health services. To test this, we implemented DID regressions similar in structure to Equation (1) using a variety of outcome variables of interest. The results of these regressions are presented in Table 3.

TABLE 3 — EFFECT ON OBJECTIVE MEASURES OF HEALTH

Regressand	PCS‡	MCS‡	Hospital visit	Doctor visit	Days unable to work
DID coefficient	−0.131	0.057	0.053	0.010	0.113
	[0.260]	[0.342]	[0.092]	[0.055]	[0.097]
(χ^2-) F-statistic	238.19	44.80	(4,490.97)	(2,442.12)	(3,507.39)
(Pseudo) R^2	0.317	0.073	(0.086)	(0.047)	—
Observations	25,646	25,646	89,897	89,897	89,897
Model†	OLS	OLS	Logit	Logit	ZINB

NOTES: Bracketed numbers are standard errors clustered at the individual level. All regressions include a full set of control variables. † We used ordinary least squares, logit, and zero-inflated negative binomal models as appropriate. ‡ The Physical Component Summary and Mental Component Summary scales are only available for the years 2002 and 2006. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
SOURCE: Own computation based on SOEP 2000–2003 and 2005–2007.

Unfortunately, the dataset does not contain an actual measure of objective health status such as those that could be obtained from biomarkers. However, we assume that the following variables, taken together, serve as a better proxy for objective health status since these could less likely be influenced by non-health-related factors. These objective measures include the Physical Component Summary (PCS) scale, Mental Component Summary (MCS) scale, the probability of staying in a hospital, the probability of visiting a doctor, and the number of days absent from work. Two studies have already shown that the reform had no effect on the probability of visiting a doctor [Augurzky, Bauer and Schaffner 2006; Schreyögg and Grabka 2008]. We confirm their results here. For the PCS and MCS scales, least-squares regressions are estimated. A logit model is used for the probability of staying in a hospital and a zero-inflated negative binomial model is used for the number of days absent from work.

Using these five additional objective measures of health, we found that the reform had no significant effect. It is therefore much more plausible that people responded to the reform by adjusting their preferences or expectations with regard to health services.

5.4 Parallel trends

The DID estimator is biased if there are divergent trends in the outcome variable for the two groups. Let $SRHS_{it,s}$ be the SRHS of individual i at time t in state s, where s is equal to 1 if the individual is exposed to treatment. The parallel-trends assumption is then:

$$\mathrm{E}\left[SRHS_{i2,0} - SRHS_{i1,0}|s = 1\right] = \mathrm{E}\left[SRHS_{i2,0} - SRHS_{i1,0}|s = 0\right], \tag{2}$$

which says that the observable $\mathrm{E}\left[SRHS_{i2,0} - SRHS_{i1,0}|s = 0\right]$ is an unbiased estimate of the unobservable counterfactual $\mathrm{E}\left[SRHS_{i2,0} - SRHS_{i1,0}|s = 1\right]$. Proceeding to estimate a DID model where Equation (2) does not hold results in an unreliable measure of the treatment effect.

Situations in which Equation (2) is not satisfied include the presence of differential secular time effects. For example, if these time effects actually depend on one's level of income, and income determines selection into the treatment group (as it does in this case with a federally mandated income threshold), then the simple DID estimator (i.e., without controls) will not properly identify the treatment effect. Therefore, we included control variables in the regression to alleviate the concern about the consistency of the estimator as well as to improve its efficiency.

In this paper, two increasingly common approaches are pursued to statistically test whether there were parallel trends in the pre-intervention periods. The first follows Di Tella and Schargrodsky [2004]. All post-intervention observations are dropped from the sample. Then, regressions are ran on the dependent variables: first, the DID variable is redefined such that the GMG became effective in 2001; second, such that the GMG became effective in 2002. We call these "placebo regressions". If the SHI and

PHI exhibited common trends before the actual intervention in 2004, then the estimated treatment effect in these placebo regressions should not be significantly different from zero. Indeed, Table 4 shows that the results from these regressions validate the DID strategy.

TABLE 4 — PLACEBO REGRESSIONS

	2001 placebo		2002 placebo	
	POL	OLS	POL	OLS
DID coefficient	0.076	0.030	0.073	−0.031
	[0.070]	[0.027]	[0.060]	[0.023]
(χ^2-) F-statistic	(8,245.99)	267.67	(8,246.60)	267.74
(Pseudo) R^2	(0.109)	0.251	(0.109)	0.251
Observations	54,768		54,768	

NOTES: Bracketed numbers are standard errors clustered at the individual level. Regressions include a full set of control variables. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
SOURCE: Own computation based on SOEP 2000–2003.

The second approach pursued here follows and expands on Galiani, Gertler and Schargrodsky [2005]. Again, all post-intervention observations are deleted and the treatment-group indicator is then interacted with all the pre-intervention year dummies except for 2000, which serves as the base year. POL and OLS models are then estimated. The coefficients of the interaction terms in these models are subsequently tested in three ways: first, if each one is separately significant; second, if they are jointly significant; third, whether they are equal to each other. For both models, these coefficients are each and jointly insignificant and are equal to each other, which reinforces the identification strategy used here.

5.5 Further results

In this section, we present four further important results. First, we estimate the effect of the reform on general life satisfaction. In the dataset, this is measured using an 11-point scale from 0 to 10, with 10 being the most satisfied. We show that the positive effect of the reform is specific to SRHS and not to general life satisfaction. Second, we perform the same analysis using an alternative control group. This is important to show that the significant treatment effect estimated in Section 5.2 is not just an artifact of using the privately insured as the control group. Third, to the extent that only those people who actually visited a doctor would change their subjective assessments of their own health, we kept only observations with at least one doctor visit in the last month and then performed the regressions again. Finally, to get an understanding of the heterogeneous impact of the reform, we split the sample into males and females to show that sex may play a role.

We found that the publicly insured experienced a reduction in their general life satisfaction. The DID-coefficient estimates [standard errors] are -0.071 [0.045] and -0.081 [0.038] for the POL and OLS models, respectively, using a full set of control variables. Both SRHS and general life satisfaction are measures of subjective well-being and in this case, we might expect them to move in the same direction. For example, a shift in the optimism of people might move these self-assessments upwards. However, we show that the reform of the health insurance system in Germany only increased self-assessed health and hence rule out confounders similar to one's optimism. If anything, the publicly insured saw a deterioration in other aspects of their lives that dominate the improvement in SRHS. This may be caused by the added burden of increased prices for health care services. That we still found a positive effect on SRHS is therefore all the more remarkable.

The GMG of 2004 provides an alternative control group because the reform only applied to the statutorily insured who are at least 18 years old. This allows us to test the robustness of our result by re-estimating the model using an alternative control group that consists only of the privately insured 18-year olds and all the 17-year olds. For this analysis, we restrict the sample to adolescents aged 17 to 18, resulting in 1,391 person–year observations, of which 461 (33 percent) constitute the control group.

We implemented a DID regression model as in Equation (1). However, the control variables now only consist of the following: sex, years of education, (log) household net income, (log) household labor income, civil status (whether married or not), employment status (indicator variables for full-time employment, part-time employment, self-employment, and not currently working), whether the respondent had a child recently, whether the respondent is disabled and the degree of disability, whether the respondent recently visited the hospital, whether the respondent recently visited a doctor, and whether the respondent is a smoker.

The regression results from the POL and OLS models support the findings in Section 5.2. The estimated DID coefficient [s.e.] from the POL model is 0.525 [0.233], while the corresponding value from OLS is 0.214 [0.092]. Notably, the estimated ATET using the alternative control group is substantially higher than that from using the pool of the privately insured as the control group. This could either mean that younger people have more adaptive preferences or are more swayed by the informational effect of prices.[19]

It is likely that only people who have actually visited a doctor would revise their preferences since they would have had to experience co-paying for the doctor visit (barring those few who were exempted). To this end, we excluded from the sample those observations for which there were no visit to the doctor in the quarter prior to the interview. Using the models with control variables, the estimated DID coefficients [s.e.'s] from the POL and OLS models are 0.128 [0.064] and 0.049 [0.027], respectively.

Finally, we note that the effect is driven by the males in the sample. The parameter of interest turns insignificant if we restrict the sample to females. However, it does carry the same sign and remains significant when the sample is restricted to males. Specifically, the estimated parameters [s.e.'s] in the models with a full set of control variables are 0.119 [0.062] in the POL model and 0.040 [0.024] in the OLS model.

[19] We confirm this by re-estimating the original model but limiting the observations only for those people who are younger than 25. We obtain a positive and significant estimate of the treatment effect in this regression. Additionally, we perform placebo regressions similar to the ones described in Section 5.4 to support the assumption of parallel trends. The results confirm that both the treatment and alternative control groups followed the same trend before the intervention.

6 Conclusion

Economists now routinely use measures of subjective well-being in their research. However, it is not clear at the outset that we can contribute much to a better understanding of these measures. As Hamermesh [2001] puts it, "[W]e have not done much to justify our incursions." Nevertheless, among the numerous subjective measures at our disposal, SRHS is among the most prevalent as an object of analysis. Its advantage is that it is highly correlated with objective measures of health. In the absence of the latter, researchers resort to its subjective counterpart. It is therefore important to understand how this subjective measure behaves in response to stimulus changes in an economic agent's environment.

In 2004, Germany introduced a major reform in its health insurance system, which effectively raised the price of health care services. Employing data from the SOEP covering the periods 2000–2003 and 2005–2007, the natural experiment resulting from this reform package is exploited to examine its impact on the SRHS of the statutorily insured. The results of a DID estimation strategy indicate that the statutorily insured rated themselves as healthier after the reform took effect. However, there is no evidence to suggest that this improved self-perception of health is a result of an actual improvement in the health status of the statutorily insured.

To explain this phenomenon, we appeal to the literature on cognitive dissonance and the informational effect of prices. People seem to have a taste for consistency and this may manifest itself when people change their actions in the present to make their past choices conform to present circumstances. This involves a certain sense of "self-deception" when people convince themselves that their optimal action *ex ante* remains optimal *ex post*. Without such self-deception, people will have to live with the psychological cost of a choice that turned out to be suboptimal *ex post* [Eyster 2002]. With respect to the informational effect of prices, there is some evidence that people judge quality by price. A higher price can induce people to think that a good is better simply because of its higher price, i.e., people value goods not merely by its price but for its price as well [Scitovszky 1944]. Unfortunately, we cannot identify which of these two mechanisms actually was in play but suffice it to say that the evidence indicates a disjoint between subjective and objective measures of health.

These subjective measures are liable to what Hamermesh [2001] calls the "scaling problem". Essentially, subjective measures are the result of a conversion of an underlying continuous ordinal measure into a measure that has only a few categories. The psychological process of this conversion from an unobservable continuous measure into a self-reported polychotomous measure is something which we as economists know very little about. Yet, national governments are keen to use these subjective measures to complement traditional indicators of economic development. For example, the government of France has sponsored and endorsed the report of the Commission on the Measurement of Economic Performance and Social Progress, which recommends, among others, using measures of subjective well-being. Therefore, it is all the more important now to look at the reliability of these measures. At least in the realm of health, we advocate the collection of more objective measures, particularly biomarkers, since we feel that these measures are less subject to the vagaries of an individual's perception of oneself.

***

References

Akerlof, George A. 1970. "The market for 'lemons': quality uncertainty and the market mechanism." *Quarterly Journal of Economics* 84(3):488–500.

Ashraf, Nava, James Berry and Jesse M. Shapiro. 2007. "Can higher prices stimulate product use? Evidence from a field experiment in Zambia." NBER Working Paper No. 13247.

Augurzky, Boris, Thomas K. Bauer and Sandra Schaffner. 2006. "Copayments in the German health system: does it work?" IZA Discussion Paper No. 2290.

Bago d'Uva, Teresa, Eddy van Doorslaer, Maarten Lindeboom and Owen O'Donnell. 2008. "Does reporting heterogeneity bias the measurement of health disparities?" *Health Economics* 17(3):351–375.

Bago d'Uva, Teresa, Owen O'Donnell and Eddy van Doorslaer. 2008. "Differential health reporting by education level and its impact on the measurement of health inequalities among older Europeans." *International Journal of Epidemiology* 37(6):1375–1383.

Bagwell, Kyle and Michael H. Riordan. 1991. "High and declining prices signal product quality." *American Economic Review* 81(1):224–239.

Bertrand, Marianne and Sendhil Mullainathan. 2001. "Do people mean what they say? Implications for subjective survey data." *American Economic Review* 91(2):67–72.

Böckerman, Petri and Pekka Ilmakunnas. 2009. "Unemployment and self-assessed health: evidence from panel data." *Health Economics* 18(2):161–179.

Bound, John. 1991. "Self-reported versus objective measures of health in retirement models." *Journal of Human Resources* 26(1):106–138.

Chamberlain, Gary. 1980. "Analysis of covariance with qualitative data." *Review of Economic Studies* 47(1):225–238.

Crossley, Thomas F. and Steven Kennedy. 2002. "The reliability of self-assessed health status." *Journal of Health Economics* 21(4):643–658.

Di Tella, Rafael and Ernesto Schargrodsky. 2004. "Do police reduce crime? Estimates using the allocation of police forces after a terrorist attack." *American Economic Review* 94(1):115–133.

Easterlin, Richard A. 1995. "Will raising the incomes of all increase the happiness of all?" *Journal of Economic Behavior and Organization* 27(1):35–47.

Eyster, Erik. 2002. "Rationalizing the past: a taste for consistency." Nuffield College, Oxford.

Ferrer-i-Carbonell, Ada and Paul Frijters. 2004. "How important is the methodology for the estimates of the determinants of happiness?" *Economic Journal* 114(497):641–659.

Frey, Bruno S. and Alois Stutzer. 2002. "What can economists learn from happiness research?" *Journal of Economic Literature* 40(2):402–435.

Galiani, Sebastian, Paul Gertler and Ernesto Schargrodsky. 2005. "Water for life: the impact of the privatization of water services on child mortality." *Journal of Political Economy* 113(1):83–120.

Gergaud, Olivier and Florine Livat. 2007. "How do consumers use signals to assess quality?" American Association of Wine Economists Working Paper No. 37296.

Haisken-DeNew, John P. and Markus Hahn. 2006. "PanelWhiz: a flexible modularized Stata interface for accessing large-scale panel data sets.".

Hamermesh, Daniel S. 2001. "The changing distribution of job satisfaction." *Journal of Human Resources* 36(1):1–30.

Hamermesh, Daniel S. 2004. "Subjective outcomes in economics." *Southern Economic Journal* 71(1):1–11.

Heckman, James, Neil Hohmann, Jeffrey Smith and Michael Khoo. 2000. "Substitution and dropout bias in social experiments: a study of an influential social experiment." *Quarterly Journal of Economics* 115(2):651–694.

Hwang, Yun-Jae, Brian Roe and Mario F. Teisl. 2006. "Does price signal quality? Strategic implications of price as a signal of quality for the case of genetically modified food." *International Food and Agribusiness Management Review* 9(1).

Kassenböhmer, Sonja C. and John P. Haisken-DeNew. 2009. "You're fired! The causal negative effect of unemployment on life satisfaction." *Economic Journal* 119(536):448–462.

Leibenstein, H. 1950. "Bandwagon, snob, and Veblen effects in the theory of consumers' demand." *Quarterly Journal of Economics* 64(2):183–207.

Mossey, Jana M. and Evelyn Shapiro. 1982. "Self-rated health: a predictor of mortality among the elderly." *American Journal of Public Health* 72(8):800–808.

Puhani, Patrick A. 2008. "The treatment effect, the cross difference, and the interaction term in nonlinear 'difference-in-differences' models." IZA Discussion Paper No. 3478.

Schreyögg, Jonas and Markus M. Grabka. 2008. "Copayments for ambulatory care in Germany: a natural experiment using a difference-in-differences approach." SOEPpapers No. 96.

Schwierz, Christoph. 2009. "Expansion in markets with decreasing demand: for-profits in the German hospital industry." Ruhr Economic Papers No. 106.

Scitovszky, Tibor. 1944. “Some consequences of the habit of judging quality by price.” *Review of Economic Studies* 12(32):100–105.

Shiv, Baba, Ziv Carmon and Dan Ariely. 2005. “Placebo effects of marketing actions: consumers may get what they pay for.” *Journal of Marketing Research* 42(4):383–393.

Wagner, Gert G., Joachim R. Frick and Jürgen Schupp. 2007. “The German Socio-Economic Panel Study (SOEP): scope, evolution, and enhancements.” *SOEPpapers* (1).

Winkelmann, Liliana and Rainer Winkelmann. 1998. “Why are the unemployed so unhappy? Evidence from panel data.” *Economica* 65(257):1–15.

Yariv, Leeat. 2005. “I’ll see it when I believe it: a simple model of cognitive consistency.” Unpublished manuscript.