G Model
JHE-1541; No. of Pages 12

ARTICLE IN PRESS

Journal of Health Economics xxx (2011) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Journal of Health Economics

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

Supply-side and demand-side cost sharing in deregulated social health insurance: Which is more effective?☆

Maria Trottmann[a,*], Peter Zweifel[a], Konstantin Beck[a,b]

[a] University of Zurich, Department of Economics, Hottingerstrasse 10, CH-8032 Zurich, Switzerland
[b] CSS Institute for Empirical Health Economics, Tribschenstrasse 21, 6002 Lucerne, Switzerland

ARTICLE INFO

Article history:
Received 1 October 2009
Received in revised form 3 October 2011
Accepted 11 October 2011
Available online xxx

JEL classification:
I13
G22
D82

Keywords:
Health insurance
Moral hazard
Managed care
Copayment
Two-stage residual inclusion

ABSTRACT

Microeconomic theory predicts that if patients are fully insured and providers are paid fee-for-service, utilization of medical services exceeds the efficient level ('moral hazard effect'). In Switzerland, both demand-side and supply-side cost sharing have been introduced to mitigate this problem. Analyzing a panel dataset of about 160,000 adults, we find both types of cost sharing to be effective in curtailing the use of medical services. However, when moral hazard mitigation is traded off against risk selection, the minimum-deductible, supply-side cost sharing option ranks first, followed by the medium-deductible demand-side alternative, making the supply-side option somewhat more effective.

1. Introduction

One of the main goals of health care financing systems is to promote efficient levels and types of care (Ellis and McGuire, 1993). If patients are fully insured and providers are paid fee-for-service, they desire larger than optimal quantities of health care services, connoted 'moral hazard'. Zeckhauser (1970) and Zweifel and Manning (2000) have analyzed how demand-side cost sharing (in the guise of deductibles or co-payments) can be used as a corrective. However, demand-side cost sharing exposes consumers to financial risk, contradicting the very objective of insurance. Unless limited by a stop-loss, it also makes beneficial procedures unaffordable to some patients (Nyman, 1999). In addition, it might be considered unfair towards the chronically ill.

These considerations have created interest in the alternative of supply-side cost sharing (in the guise of capitation or prospective payment). Because of their information advantage, providers of medical care can influence the demand for their services to a greater extent than other professionals (Arrow, 1963). Moreover, providers are less vulnerable to risk than patients because they can pool treatment cases. However, supply-side cost sharing might also promote a reduction in quality or the denial of beneficial but costly services, a phenomenon commonly termed stinting (Newhouse, 2002).

Both demand-side and supply-side cost sharing have been empirically examined in terms of their effectiveness. The novelty of this paper is that it directly compares the expenditure effects of demand-side and supply-side cost sharing (and combinations thereof), using contract variants offered by the same health insurer. This has the advantage that many side conditions (underwriting policy, billing procedure) are kept constant. Moreover, the paper complements Lehmann and Zweifel (2004), who construct a proxy for unobserved health status from prior health care expenditure (HCE), by the two-stage residual inclusion estimator (2SRI, Terza et al., 2008). In this way, risk-selection effects are more fully controlled for. Finally, it extends the set of instruments influencing choice of plan but not HCE by including the premium for the baseline contract, the potential premium reduction for a restricted plan, the individual's credit record, and years of membership with the

☆ We would like to thank the CSS Institute for Empirical Health Economics for providing data, financial support, and many insights. We are grateful to three anonymous reviewers as well as Randall Ellis, Wynand Van de Ven, Hansjoerg Lehmann, Johannes Schoder, Stefan Pichler, Lorens Helmchen, Matthias Kifmann, Ute Kunze, Ute Studer-Merkle, Bernhard Keller, Stefan von Rotz, Urs Kaeser, and participants of the 2011 Darmstadt Conference on Inequality, Insurance and Health for helpful comments and discussion.
* Corresponding author. Tel.: +41 44 634 45 96.
E-mail address: maria.trottmann@econ.uzh.ch (M. Trottmann).

0167-6296/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.jhealeco.2011.10.004

same fund. For the capitated plan, an additional instrument is a dummy indicating whether or not an Independent Practice Association (IPA) was operative in the individual's county of residence.

The data come from Switzerland, a country where consumers have annual free choice of plan with no employer involvement. The chronically ill are not precluded from switching due to open enrollment. Low-income individuals (about 30 percent of the population) are eligible for premium subsidies. Receiving the subsidy, they are less likely to choose high-deductible plans, because a reduction in the income transfer to the sick state is particularly disadvantageous if the income effect on medical consumption is strong (Nyman, 1999). On the other hand, managed-care type plans that are too restrictive compared to premium charged will not be chosen.

The remainder of this article is structured as follows. Section 2 contains an overview of the empirical literature. The policy setting is described in Section 3, while Section 4 is devoted to a description of the data base. In Section 5, we explain the econometric methods used to separate moral hazard from risk-selection effects and to deal with the very skewed distribution of the HCE data. The estimation results are presented in Section 6. Section 7 discusses policy implications in view of related literature, while the final Section 8 contains a summary and conclusions.

2. Literature review

In order to keep this review concise, we focus on empirical papers that measure moral hazard in health insurance. When individuals have a choice of plan, risk-selection effects need to be accounted for because those who expect high future HCE are more likely to opt for more comprehensive insurance. A small number of researchers have avoided this selection problem by benefiting from randomized experiments (the famous RAND study; Manning et al., 1987) or natural experiments (Chiappori et al., 1998; Eichner, 1998; Winkelmann, 2004). Other papers have used econometric techniques to address endogenous plan choice. Many econometric approaches require for identification the availability of at least one variable that influences contract choice but not utilization (an 'identifying instrument'). Pertinent studies from Switzerland are Schellhorn (2001), Gerfin and Schellhorn (2006) and Gardiol et al. (2006). The former two rely on premium level and supplementary hospital insurance as identifying instruments, while the latter uses death as an indicator of morbidity which is unaffected by insurance. Using Australian data, Cameron et al. (1988) advocate income as determinant of insurance coverage but not utilization. In the United States, employers play a strong role in determining the individual's choice of plan, making their characteristics potential identifying instruments. For example, Dowd et al. (1991) and Cardon and Hendel (2001) exploit the fact that different employers offer different premiums and copayment levels, while Deb and Trivedi (2009) use the employer's type (public or private), the size of the firm, and whether or not it offers both HMO and non-HMO options.

Turning to estimation techniques, one notices that instrumental variable estimators are rarely applied to non-linear frameworks. An early exception is Dowd et al. (1991), who estimate a Tobit model with a correction for selectivity (Lee, 1978). In addition, Deb and Trivedi (2009) and Deb et al. (2006) specify a fully parametric model of both choice and utilization equations, which is jointly estimated by maximum simulated likelihood. However, these approaches depend upon restrictive distributional assumptions. As HCE data are very skewed and the distribution of the 'tail' is difficult to specify correctly, Terza et al. (2008) advocate the two-stage residual inclusion estimator. It yields consistent estimates over a wide range of non-linear specifications.

Studies that have addressed endogeneity in non-linear panel data models are even more rare. Non-linear fixed-effects models are plagued by the incidental parameters problem (see Lancaster (2000) for an overview, and Chamberlain (1980) for a corrective). In random effects specifications, the incidental parameters problem can be avoided by integrating out the individual-specific effects (Vella and Verbeek, 1998, 1999). However, this requires a parametric specification of their distribution.

An alternate approach of exploiting the information of panel data was pioneered by Wolfe and Godderies (1991). It uses HCE from prior years to proxy unobserved differences between individuals which become predetermined in the year when the comparison between plans is performed (Lehmann and Zweifel, 2004; Van Kleef et al., 2008). In this paper, a combination of the IV and the 'health proxy' approach will be applied.

3. Swiss health insurance

Swiss health insurance is of the 'managed competition' type (see Kreier and Zweifel (2010) for a comprehensive description). Coverage is mandatory for a rather comprehensive 'basic' basket of medical services and pharmaceuticals, written by some 80 private, not-for-profit insurers competing in a regulated market. Free consumer choice of plan is a distinctive feature of the system. There is no pre-selection of plans by employers or government agencies. Insurers are obliged to accept all applicants during annual open enrollment periods. Premium subsidies for low-income individuals are funded out of general taxation. Premiums can be differentiated by area of residence but not by health risk. Reductions are possible for young adults (19–25) and individuals who receive accident coverage through the employer.

In the baseline contract, insured individuals enjoy unlimited access to all licensed physicians and most hospitals in their region of residence. They face a minimum annual deductible of CHF 300 (some EUR 200 as of 2006) and a copayment rate of 10 percent up to a cap of CHF 700 (EUR 470) per year. Physicians in independent practice are reimbursed fee-for-service (FFS) according to an administered fee schedule that is collectively bargained between the providers' and the insurers' associations. Hospitals receive per diems for patients treated (the nation-wide introduction of a DRG system is scheduled for 2012). The cantons[1] finance hospital investment and one-half of operational cost. While this system is generally found to ensure access to comprehensive health care to all citizens, it is criticized for high and rapidly increasing HCE, lack of co-ordination between providers, and lack of information about quality and efficiency (OECD, 2006).

In response to these problems, insurers have been granted the right to offer managed-care type options (since 1994) and higher deductibles (since 1996) in return for lower premiums. However, policy makers feared that these options would attract low risks. In addition to a risk adjustment scheme based on age and sex, they imposed limits on possible premium reductions. For voluntary deductibles, these are fixed percentages of the base premium or 80 percent of the additional financial risk taken by the consumer (deductible minus 300), whichever is less. The eligible deductible levels are also regulated, as shown in Table 1. In managed-care type contracts, the insurer must prove that the reduction is justified by efficiency gains rather than risk-selection effects. Furthermore, it must not exceed 20 percent during the first five years since

[1] Switzerland is divided into 26 cantons, with population ranging from 1,307,600 (Zurich) to 15,500 (Appenzell i.R.), Source: Swiss Federal Statistical Office, http://www.bfs.admin.ch.

Table 1
Regulation of deductibles and maximum premium reductions, 2006.

Deductible level in CHF/year	300	500	1000	1500	2000	2500
Max reduction in percent of the base premium	–	5	15	30	38	43
Max absolute reduction: 0.8 * (Deductible – 300)	–	160	560	960	1360	1760

CHF 1 ≈ EUR 0.66.

Table 2
Descriptive statistics according to type of contract, 2006.

Contract	*N*	Age	Male	HCE	Share with HCE > 0	HCE if >0
A. FFS plans						
Minimum DED (baseline contract)	84,053	55	0.4	4610	0.88	5230
		(18)	(0.49)	(8961)	(0.33)	(9378)
Medium DED	31,573	54	0.43	3229	0.81	3908
		(16)	(0.49)	(6962)	(0.38)	(7522)
High DED	38,386	45	0.54	1057	0.57	1804
		(13)	(0.5)	(3593)	(0.49)	(4580)
B. IPA plans						
Minimum DED	4942	54	0.44	2933	0.85	3427
		(17)	(0.5)	(5999)	(0.35)	(6355)
Medium DED	1134	49	0.45	1686	0.78	2121
		(15)	(0.45)	(3654)	(0.41)	(3999)
High DED	3598	43	0.56	834	0.58	1415
		(12)	(0.56)	(3431)	(0.49)	(4368)

Standard deviations in parentheses, *N* = 163,686, DED = deductible, IPA = Independent Practice Association, CHF 1 ≈ EUR 0.66.

the launch of the contract. The same deductible levels apply to managed-care type and FFS plans.

4. Data

The data base consists of individual records of more than 160,000 Swiss adults insured by CSS, a major Swiss insurer, and covering the years 2003–2006. It includes age, gender, residential location, contract choice, and HCE. Individuals who were not observed over the entire four years are excluded from the analysis, with death constituting the main cause. While the deathbound are known to cause a considerable amount of HCE, they exhibit an idiosyncratic pattern of health care utilization (see Werblow et al., 2007), justifying separate analysis. The influence of closeness to death also calls for exclusion of individuals who died during 2007, resulting in a panel comprising some 160,000 individuals.

The supply-side cost sharing variant analyzed here is an IPA. Similar to the United States, participating physicians (mainly general practitioners) are paid a risk-adjusted capitation payment designed to cover all services rendered or prescribed up to a threshold of CHF 10,000 per patient and year. Beyond that limit, the insurer reimburses 90 percent of cost, as calculated according to the FFS alternative. Capitation payments are adjusted for age, gender, deductible level, hospitalization during previous year, nursing home stay during previous year, and 21 pharmaceutical cost groups. The pharmaceutical cost groups are similar to those used in the Dutch risk adjustment scheme (Lamers and Van Vliet, 2003). While the insurer does not impose guidelines or utilization reviews, many networks run them internally, combined with quality monitoring by independent auditors in some cases.

Table 2 shows descriptive statistics according to contract choice as of 2006. For simplicity, deductibles are grouped into three categories (minimum: 300, medium: 500, high: ≥1000 CHF per annum). More than 70 percent of those who chose a deductible in excess of CHF 500 opted for the CHF 1500 level. The other high deductible levels were only chosen by a small number of individuals each, so including them separately would have resulted in unstable estimates. Furthermore, observed HCE across the high deductible levels appeared to be similar.[2]

According to panel A of Table 2, buyers of the high-deductible FFS plans are younger and more likely to be male than those with a minimum deductible. Their mean HCE amounts to CHF 1057 or 23 percent of the CHF 4610 pertaining to individuals with the minimum deductible. Also, their fraction of reporting positive HCE is 57 rather than 88 percent. If only those with positive HCE are taken into account, the mean is CHF 1804 or 34 percent of the minimum-deductible benchmark of CHF 5230, respectively. These differences point to sizeable effects of demand-side cost sharing (which still need to be corrected for risk-selection effects, see below).

Turning to the supply-side cost sharing alternative (panel B of Table 2), one notices that the HCE values for the IPA plans are lower throughout than for the conventional FFS plans with the same deductible level. In the minimum deductible group, average age is similar in the IPA and in the FFS plan. However, the high-deductible IPA variant is again characterized by a comparatively low mean age and a higher share of men. Average HCE is CHF 834 or 18 percent of the minimum-deductible, FFS benchmark of CHF 4610. The share of individuals with positive HCE is 58 percent rather than 88 percent, while mean HCE conditional on being positive amounts to CHF 1415 or 27 percent of the CHF 5230 benchmark.

In order to get a preliminary indication of the extent to which the cost differences may be caused by risk-selection effects, it is

[2] For patients with high deductibles, it is questionable how well their HCE are observed. In earlier work (e.g. Lehmann and Zweifel, 2004), only individuals with the minimum deductible were analyzed on the grounds that patients have no incentive to submit their claims unless HCE exceeds the deductible. However, with the advent of electronic billing systems, the lion's share of billings are now transmitted directly from providers to insurers, who then charge the deductible to the patient. In some cantons, physicians even decided to abandon direct-to-consumer billings completely. CSS conducted an internal study relating the share of direct-to-consumer billing to HCE below the deductible. Contrary to expectations, the billing mode interacted with the deductible level had no influence on the probability of reporting positive HCE. It also had no significant impact on explaining positive HCE. It appears that many individuals submit their bills regardless of their deductible, maybe to decrease the administrative burden in case of an illness. In the dataset, 24 percent of individuals with positive HCE had HCE below their deductible.

Table 3
Prior-year mean HCE of switchers and non-switchers.

Switch at the start of 2006, HCE in 2005	FFS, minimum DED to medium DED	FFS, minimum DED to high DED	FFS (all DED) to capitated IPA
Non-switchers	4315	4315	3230
Switchers	1916	826	1597

DED = deductible, IPA = Independent Practice Association, FFS = fee-for-service, CHF 1 ≈ EUR 0.66.

instructive to compare the HCE of consumers who switch to higher deductibles and IPA plans with the HCE of those who do not (see Table 3). The switchers from a minimum-deductible FFS plan in 2005 to a high-deductible one in 2006 had caused HCE of CHF 826, a mere 19 percent of the non-switchers. Those changing to a medium-deductible alternative had caused HCE amounting to CHF 1916 in 2005, or 44 percent of the non-switchers. Switchers who moved from a FFS to an IPA plan had HCE amounting to CHF 1597, or 49 percent of the stayers. These figures point to substantial risk-selection effects in both demand-side and supply-side cost sharing.

5. Econometric model

5.1. Developing a proxy for unobserved health status, 2003–2005

The dataset does not contain direct information on health status such as diagnostic codes, restrictions on activities of daily living, or self-reported health. However, panel data allows to develop an indicator of health status from prior HCE (see Van Kleef et al., 2008; Lehmann and Zweifel, 2004; Wolfe and Godderies, 1991). In particular, Lehmann and Zweifel show how residuals from a random-effects Tobit regression of prior HCE on exogenous variables can serve as a proxy for unobserved health.

However, in view of considerable heteroscedasticity in the dataset, the two-part model is preferable over the Tobit. The first part is a random-effects probit model predicting the probability of observing positive HCE for individual i in year t [see Eq. (1)]. The second part estimates the amount of HCE given that it is positive. The log transformation serves to reduce the skewness of the dependent variable. The present panel is unbalanced, as many individuals had positive HCE in some but not in all years. A Wooldridge test of serial correlation in the error term (Wooldridge, 2002) rejected the null hypothesis of no autocorrelation. Therefore, the feasible generalized least squares procedure proposed for unbalanced panels by Baltagi and Wu (1999) is applied to gain efficiency while avoiding biased estimation of standard errors. The model for deriving the health status proxy thus reads (all error terms assumed normally distributed),

$$Pr(\mathrm{HCE}_{i,t} > 0) = \Phi(a + \beta X_{i,t} + \alpha_i + \upsilon_{i,t}) \quad (1)$$

$$\log(HCE_{i,t} | HCE_{i,t} > 0) = b + \theta X_{i,t} + \gamma_i + \epsilon_{i,t} \quad (2)$$

with $\epsilon_{i,t} = \rho * \epsilon_{i,t-1} + \xi_{i,t}$ Eqs. (1) and (2) are estimated on the first three years of the dataset, i.e. 2003–2005. Explanatory variables are age, age interacted with gender, urbanization, area of residence, and a year dummy to account for inflation. Estimation results are shown in Table A.10 of the appendix. Deviations from the expected value of HCE are averaged over the three years in order to reduce the influence of transitory health shocks.[3]

5.2. Endogeneity of contract choice, 2006

Even if the proxy derived from Eqs. (1) and (2) controls for unobserved differences in current health status, there are additional unmeasured variables that may cause someone opting for the minimum deductible to have a great deal of HCE, resulting in an overestimation of moral hazard effects. Examples are private information about probabilities of future illness, general attitude towards medical care, and previous experience with the health care system. Ignoring these confounders will lead to omitted variable bias in the HCE equation. Terza et al. (2008) show that the residuals from an equation modeling contract choice are good estimators of these confounders. Therefore, these residuals are included in the HCE equation alongside observed contract choice and the proxy for latent health (two-stage residual inclusion estimation, 2SRI). The 2SRI method also yields consistent estimates if the HCE equation is nonlinear. However, it requires equations for contract choice to be specified.

For identification, at least one explanatory variable in the contract choice equation must not appear in the HCE equation. Five such variables are available.

1. Baseline premium[4]: A high baseline premium increases the attractiveness of higher-deductible and IPA options. At the same time, there is little reason why premiums should influence health care consumption. Their income effect is limited in the Swiss case because low-income individuals (some 30 percent of the population) are eligible for a premium subsidy. Moreover, preliminary estimations showed that premiums do not influence HCE when other factors are controlled for.[5]
2. Absolute premium reductions for a higher-deductible or an IPA option: While premium reductions make these contract options more attractive, they should not influence health care consumption for the same reasons as described in item no. 1 above. They were also found to be insignificant in an estimation of HCE.
3. Number of years of CSS membership[6]: Long-standing members are known not to switch contracts, making them less likely to opt for a higher-deductible or an IPA option. However, loyalty is negatively correlated with health status because consumers who develop chronic conditions face a premium hike if they sign up with another insurer for the supplementary component (which they usually prefer to have from the same insurer to avoid ambiguity as to responsibility for payment). Nevertheless, preliminary estimations showed it to be insignificant in the HCE equation when entered in combination with the health status proxy. It therefore qualifies as an identifying restriction.
4. Dummy indicating a bad credit record: This may reflect lower income, which is relevant for contract choice because high

[3] Note that while only individuals with the baseline contract are included here, estimated coefficients will be used to predict individual HCE for the whole sample. This has the advantage that the endogeneity of contract choice does not bias estimators. In order to test robustness, we also calculated the proxy including all individuals, with little effect on results.

[4] Although we use data from only one insurer, premiums differ between regions. In addition, young persons and individuals who have accident coverage through their employer are eligible for premium reductions.

[5] Premium levels were also used as identifying instruments in Schellhorn (2001).

[6] This variable is truncated at 1999 because retrieving data from earlier years is cumbersome. There was a change in IT architecture in 1998.

deductibles are unattractive to risk-averse low-income individuals. At the same time, a bad credit record proved unrelated to HCE once the proxy for health status was included.
5. IPA officially on offer within the individual's area of residence: The availability of an IPA importantly favors the choice of the corresponding option. However, it proved to be unrelated to utilization provided regional differences were controlled for by dummies.

Modeling the choice of deductible calls for an ordered probit model, while for the choice of the IPA a probit model is sufficient. For the probit, the generalized residuals were derived by Gourieroux et al. (1987). Let h_i be an indicator variable equal to one if the IPA plan was chosen and zero otherwise, z_i a vector of covariates, and $\hat{\theta}$ a vector of the estimated coefficients. Then, the generalized residuals $\hat{u}_i$ are given by

$$\hat{u}_i = h_i * \frac{\phi(z_i'\hat{\theta})}{\Phi(z_i'\hat{\theta})} + [1-h_i] * \frac{-\phi(z_i'\hat{\theta})}{1-\Phi(z_i'\hat{\theta})} = \frac{[h_i - \Phi(z_i'\hat{\theta})]\phi(z_i'\hat{\theta})}{[1-\Phi(z_i'\hat{\theta})]\Phi(z_i'\hat{\theta})}, \quad (3)$$

where Φ denotes the cumulative and ϕ, the standard normal density, respectively. In the same spirit, the generalized residuals for multinomial or ordered choice models have been defined by Vella (1993). Let there be $i = 1, \ldots, N$ individuals choosing from $k = 1, \ldots, K$ ordered alternatives, and let d_{ik} denote an indicator function taking the value 1 if individual i has chosen alternative k and zero otherwise. Then, generalized residuals $\hat{v}_i$ are given by

$$\hat{v}_i = \Sigma_{k=1}^{K} d_{ik} \frac{\hat{\pi}_{ik}[d_{ik} - \hat{\Pi}_{ik}]}{[1-\hat{\Pi}_{ik}]\hat{\Pi}_{ik}} \quad (4)$$

with $\hat{\Pi}_{ik}$ denoting the estimated cumulative probability that individual i chooses the kth alternative and $\hat{\pi}_{ik}$, the estimated value of the density at that point. These two quantities are determined as follows. Let $\hat{\gamma}$ be the vector of estimated coefficients from the ordered probit and $\hat{\alpha_k}$, the estimated cut points with $\alpha_0 = -\infty$ and $\alpha_K = \infty$. Then,

$$\hat{\pi}_{ik} = \phi(\alpha_{k-1} - z_i'\hat{\gamma}) - \phi(\alpha_k - z_i'\hat{\gamma}) \quad \textit{and} \quad \hat{\Pi}_{ik} = \Phi(\alpha_k - z_i'\hat{\gamma}) - \Phi(\alpha_{k-1} - z_i'\hat{\gamma}). \quad (5)$$

5.3. Specification of the two-part model, 2006

The distribution of HCE has a cumulation point at zero. Among the alternatives available for dealing with this fact, the two-part model is preferred over e.g. the Tobit model for two reasons. First, the zeroes are perceived as reflecting choices rather than missing values (see Jones, 2000). Second, both supply- and demand-side cost sharing are known to affect the decision to use health care at all differently from the decision how much care to use.

The first part of the two-part model is often specified as a probit and estimated by maximum likelihood. However, with the inclusion of the residuals from the contract choice equations, its errors are non-normally distributed, causing maximum likelihood to be inconsistent.[7] To avoid this problem, a GLM estimation with a probit link is applied to the HCE data of 2006. This method is consistent as long as $E(l|x) = \Phi(x\beta)$, with l denoting an indicator that equals one if HCE > 0 and zero otherwise. Normal distribution of the error terms is not required (Cameron and Trivedi, 2005).[8]

The specification of the second part (HCE|HCE > 0) has been discussed by Manning (1998) and Manning and Mullahy (2001) (MM hereafter). Because of the positive skewness of the dependent variable, raw-scale estimates can be imprecise even in large datasets. The log transformation is often used to mitigate skewness, with coefficients interpreted as (semi-) elasticities of the mean response. However, Manning (1998) shows that if the error variance is heteroscedastic in a way that is correlated with the covariates, these coefficients are no longer consistent elasticity estimates. Moreover, a retransformation of predicted values is required as soon as absolute rather than relative savings due to cost sharing are of interest.

Blough et al. (1999), MM and others suggest estimating ln ($E[y|x, y > 0]$) directly by a GLM procedure with a log link (i.e. $\ln(E[y|x, y > 0]) = x\beta$) and an appropriate variance function. As MM point out, the GLM estimates are consistent as long as the mean function is correctly specified, but might lead to imprecise estimates if the residuals are positively skewed even after transformation to log. Following the procedure for model selection suggested by MM, we start with a consistent GLM procedure, the gamma regression. The kurtosis of the 2006 residuals on the log scale is 3.53. This creates a tradeoff between imprecision (GLM) and possible bias (OLS applied to ln(y)). In this work, GLM is used because taking heteroscedasticity into account is deemed more important than precision. Given GLM, a Park test is performed to select the variance function. The estimated λ is 1.81, which is closest to the gamma specification.

The residuals from the contract choice equations are estimates rather than observations. Not accounting for this in the outcome estimation could lead to downward biased estimates of the standard errors (Heckman, 1976). Therefore, the standard errors were obtained by bootstrapping.[9]

6. Results

6.1. Effects of demand-side and supply-side cost sharing

The results for the first part of the two-part model estimated on HCE data for 2006 are shown in the first three columns of Table 4. The first column pertains to the full model. The second column excludes the residuals from the contract choice equation. The third column corresponds to a naive specification that also excludes the proxy for health status. For the variables of interest (DED, IPA), marginal effects are calculated for a representative individual (in italics below the coefficients), i.e. a woman at the age of 52, living in a suburban community in the Zurich region, having the baseline contract plus accident coverage, and a supplement covering alternative medicine. The health proxy is taken at its sample average. This individual's estimated probability of positive HCE is roughly 91 percent. For interaction terms, marginal effects are calculated according to the formulas provided by Norton et al. (2004).[10]

[7] The generalized residuals are non-linear transformations of normally distributed variables. These are not normally distributed. Then, the errors from the two-part model are linear combinations of several normally distributed variables and one non-normally distributed variable. This combination is not normally distributed.

[8] There is no contradiction to Terza et al. (2008), who suggested estimation by non-linear least squares. Although it goes by a different name, GLM is an iteratively reweighted nonlinear least squares estimator (Hardin and Hilbe, 2007).

[9] The estimation was repeated 400 times after resampling with replacement (clustered by patient). In our specific application, the bootstrapped standard errors turned out to be similar to those obtained in the original estimation.

[10] To be specific, let β_a and β_b be the coefficients of two dummies, β_{ab} the coefficient of their interaction and $\bar{x}'\beta$ the influence of all other variables at representative values. The marginal effect of the interaction term is $\Phi(\beta_a + \beta_b + \beta_{ab} + \bar{x}'\beta) - \Phi(\beta_a + \bar{x}'\beta) - \Phi(\beta_b + \bar{x}'\beta) + \Phi(\bar{x}'\beta)$. For a dummy without interaction, the marginal

Table 4
Estimation results from the two-part model, 2006.

	P(HCE>0), GLM with probit link			HCE|HCE >0, GLM with log link		
	Full	Restricted	Naive	Full	Restricted	Naive
Affluent community	0.039	0.041	0.056**	0.023	0.025	0.044
	(0.026)	(0.023)	(0.020)	(0.028)	(0.030)	(0.029)
Regional center	−0.052**	−0.052**	−0.036*	−0.092**	−0.090**	−0.084**
	(0.022)	(0.019)	(0.017)	(0.026)	(0.028)	(0.027)
Suppl. hospital	0.065***	0.069***	0.155***	0.030*	0.028	0.072***
	(0.011)	(0.011)	(0.010)	(0.014)	(0.015)	(0.014)
Suppl. altern. med.	0.036***	0.035***	0.085***	−0.076***	−0.078***	−0.063***
	(0.01)	(0.010)	(0.009)	(0.013)	(0.014)	(0.013)
Health proxy	0.473***	0.469***		0.523***	0.521***	
	(0.01)	(0.009)		(0.013)	(0.013)	
(Health proxy)^2	−0.061***	−0.061***		−0.085***	−0.085***	
	(0.004)	(0.004)		(0.007)	(0.006)	
(Health proxy)^3	0.005***	0.005***		0.003***	0.003***	
	(0.000)	(0.000)		(0.001)	(0.001)	
$\hat{v}_i$ DED	0.059			−0.332		
	(0.145)			(0.254)		
$\hat{u}_i$ IPA	−0.046***			−0.006		
	(0.013)			(0.019)		
Medium DED	−0.042**	−0.079***	−0.265***	−0.071***	−0.072***	−0.260***
	(0.016)	(0.012)	(0.010)	(0.014)	(0.016)	(0.015)
ME	*−0.007***	*−0.013****	*−0.055****			
	(0.003)	*(0.002)*	*(0.003)*			
Medium DED * IPA	0.084	0.083	0.122*	0.013	−0.001	0.053
	(0.066)	(0.061)	(0.054)	(0.071)	(0.078)	(0.075)
ME	*0.012*	*0.013*	*0.025***			
	(0.008)	*(0.008)*	*(0.011)*			
High DED	−0.207***	−0.295***	−0.869***	−0.182***	−0.191***	−0.837***
	(0.026)	(0.011)	(0.009)	(0.014)	(0.019)	(0.017)
ME	*−0.037****	*−0.054****	*−0.241****			
	(0.005)	*(0.003)*	*(0.005)*			
High DED * IPA	0.043	0.043	0.126***	0.066	0.054	0.188***
	(0.037)	(0.037)	(0.033)	(0.071)	(0.059)	(0.057)
ME	*0.011*	*0.014**	*0.034****			
	(0.006)	*(0.006)*	*(0.009)*			
IPA	0.072	0.086**	−0.066**	−0.118*	−0.189***	−0.418***
	(0.042)	(0.028)	(0.025)	(0.059)	(0.035)	(0.033)
ME	*0.011*	*0.012****	*−0.012**			
	(0.006)	*(0.004)*	*(0.005)*			
AIC	0.676	0.676	0.885	18.088	18.088	18.365
N	163,686	163,686	163,686	128,744	128,744	128,744

DED = deductible, IPA = Independent Practice Association, standard errors in parentheses, ME = estimated marginal effects. Additional regressors are age, gender, additional types of municipalities, region specific dummies, regulation on drug dispensing, accident coverage, long term care coverage, and youth rebate eligibility. $^*p<0.05$, $^{**}p<0.01$, $^{***}p<0.001$.

Voluntary deductibles are found to progressively reduce the probability of positive HCE; however, their incentive effect shrinks to 0.7 percentage points for the medium and 3.7 points for the high-deductible category, respectively, according to the full specification. Quite generally, estimated values are about two to seven times smaller in the full and restricted than in the naive specification, pointing to considerable risk-selection effects based on health status. On the other hand, they cannot be said to depend on the type of plan (DED * IPA insignificant except in the naive specification). By way of contrast, membership in an IPA may be associated with a higher probability of positive HCE (significant only in the restricted model), possibly reflecting preventive services offered. For example, one large IPA hands out vouchers for free immunizations against the flu in the fall.

Two groups of variables of Table 4 are worth a comment. First, supplementary hospital coverage (mainly for a private room) is associated with a higher probability of using medical services even after controlling for possible risk-selection effects. The same is true of supplementary coverage of therapies belonging to the category of alternative medicine. Second, the squared and the cubic form of the health proxy are highly significant as well. The impact of past, time-invariant health status on current expenditure thus does not appear to be linear over the whole distribution of HCE.

In the second part of the model, the amount of positive HCE is estimated (last three columns of Table 4). Higher voluntary deductibles are again found to progressively reduce HCE in the full and the restricted specifications, with 'true' savings due to incentive effects amounting to 7 percent (medium deductible) and about 18 percent (high-deductible category), respectively. This time, the full and restricted models point to effects that are about four times smaller than according to the naive model. Type of contract does not play a role (DED * IPA insignificant except in the naive specification). Turning to supply-side cost sharing in the guise of an IPA, one notices a reduction of HCE amounting to 12 percent (full model) and 19 percent (restricted model) respectively, less than one-half of the 42 percent suggested by the naive specification.

6.2. Estimating moral hazard and risk-selection effects

For policy, cost savings in Swiss Francs (CHF) rather than in percent are of interest. Unlike relative savings, these strongly depend

effect is $\Phi(\beta_a + \bar{x}'\beta) - \Phi(\bar{x}'\beta)$. As these marginal effects are combinations of all coefficients, their standard errors are calculated by the delta method. The calculations are run in STATA using the nlcom command.

Table 5
Estimated cost reductions in Swiss Francs, 2006.

	$\widehat{\text{HCE}}$ according to subpopulations:			
	Actual choice of contract (1)	Simulated baseline contract (2)	Moral hazard (2) – (1)	Risk selection (see text)
A. FFS plans				
Minimum DED (baseline)	4320			
	(29)			
Medium DED	3180	3430	250	891
	(37)	(34)		
High DED	1100	1422	322	2898
	(23)	(57)		
B. IPA plans				
Minimum DED	2985	3340	355	980
	(99)	(133)		
Medium DED	2023	2374	351	1946
	(159)	(124)		
High DED	843	1111	268	3209
	(55)	(78)		

Standard errors in parentheses, DED = deductible, IPA = Independent Practice Association.

on the expenditure level of the subpopulation who chooses the respective contract. The results are displayed in Table 5. For estimate (1), expected HCE according to type of contract is estimated by predicting the probability of positive HCE times the amount of HCE. For instance, individuals with the baseline contract had expected HCE of CHF 4320 (the reference value), while those with a high deductible FFS contract had CHF 1100 only. These values are derived from the full specifications displayed in Table 4, which control for both health-related and other determinants of contract choice.

Note that the value of CHF 1100 is the estimated average expenditure of the individuals *who actually chose the high-deductible contract*. In order to estimate the expenditure of the *same* subpopulation assuming they had chosen the baseline contract, the dummies for deductibles (or IPA plans, respectively) are set to zero when predicting expected HCE. The results are shown as estimate (2) of Table 5. Since both estimates (1) and (2) pertain to the same subpopulation of individuals, their difference represents the influence of moral hazard.

For instance, the effects of demand-side cost sharing can be deduced from estimate (2) for the high-deductibles subpopulation in panel A of Table 5. This is the same subpopulation that gives rise to estimate (1), with determinants of contract choice held constant. The only difference is that their predicted HCE is derived by setting the DED dummy equal to zero. Since treating physicians were confronted with the same incentives, the difference of CHF 322 (=1422 – 1100) in all likelihood is caused by the difference in demand-side incentives.

This estimate can now be compared to the effect of supply-side incentives. In panel B of Table 5, estimate (1) for the subpopulation in the minimum-deductible category amounts to CHF 2985, while estimate (2) amounts to CHF 3340. Again, both estimates refer to the same subpopulation, except that the IPA dummy is set to zero in estimates (2). Since determinants of contract choice are controlled for and the same minimum deductible is applied to the subpopulation, the difference of CHF 355 (=3340 – 2985) can be attributed to the difference in supply-side incentives.

When comparing the entries of panels A and B of Table 5, one is led to conclude that the effects of supply-side and demand-side cost sharing have about the same magnitude. However, both types of moral hazard mitigation come at the price of considerable risk-selection effects. They can be estimated as follows. In panel A of Table 5, the high-deductible subpopulation has an estimated HCE of 1100, while the subpopulation with the baseline (minimum-deductible) contract has CHF 4320. With the determinants of contract choice not held constant this time, the difference of CHF 3220 is caused by both moral-hazard and risk-selection effects. Mitigation of demand-side moral hazard has been estimated at CHF 322 above; therefore, the remainder of CHF 2898 (=3220 – 322) needs to be attributed to risk selection. Turning to supply-side cost sharing in panel B of Table 5, the high-deductible IPA subpopulation is seen to have estimated HCE of CHF 843, even CHF 3477 below the benchmark value of CHF 4320. However, only CHF 268 of this difference can be traced to an attenuation of moral hazard, leaving CHF 3209 (=3477 – 268) as the likely effect of risk selection. A comparison of the entries of panel A and B of Table 5 reveals that supply-side cost sharing in combination with demand-side cost sharing seems to go along with even more marked risk-selection effects than demand-side cost sharing combined with FFS.

Overall, the estimates of Table 5 point to a conflict of interest. If the objective is simply to reduce HCE, high-deductible plans, preferably combined with the IPA option, fare best. If however the objective is to achieve a favorable balance between moral hazard reduction and risk-selection effects, the minimum-deductible IPA option ranks first, with a ratio of 355/980, or 1:2.76, followed by the medium-deductible FFS alternative with 250/891 or 1:3.56. These ratios are of interest to policy makers or regulators who wish to introduce voluntary cost-sharing plans in order to mitigate moral hazard, but at the same time wish to avoid risk-selection effects that lead to market segmentation and hence high premiums for high risks. If the risk-selection problem is regarded as severe, plans with low ratios of moral hazard reduction versus risk selection are preferable.

6.3. Tests of validity

Two types of validity tests are performed in this section. A first set contains modifications of the two-part model in terms of observation period, estimation of the health status proxy, and econometric specification. The second type of test focuses on individuals who switched away from the minimum-deductible plans between 2003 and 2006.

In the two columns (A) of Table 6, results are shown for a pooled GLM estimation over the years 2004–2006 rather than just 2006 (standard errors are clustered by patient). In an attempt to measure variations in health status over time, the proxy now is calculated per year using data from the respective previous year.

Table 6
Specification tests.[a]

	P(HCE > 0), GLM with probit link			HCE|HCE > 0, GLM with log link			FE
Test	A	B	C	A	B	C	D
Data for 2PM	2004–06	2005–06	2006	2004–06	2005–06	2006	2004–06
Medium DED	−0.137***	−0.097***	−0.058***	−0.117***	−0.074***	−0.091***	−0.003
	(0.007)	(0.009)	(0.015)	(0.011)	(0.015)	(0.016)	(0.014)
Med. DED * IPA	0.055	0.080	0.093	−0.037	−0.047	0.020	−0.154*
	(0.043)	(0.059)	(0.062)	(0.060)	(0.081)	(0.079)	(0.071)
High DED	−0.441***	−0.329***	−0.233***	−0.265***	−0.185***	−0.218***	−0.149***
	(0.008)	(0.011)	(0.025)	(0.016)	(0.021)	(0.033)	(0.017)
High DED * IPA	0.093***	0.070*	0.062	−0.041	−0.077	0.090	−0.091
	(0.024)	(0.030)	(0.038)	(0.048)	(0.056)	(0.060)	(0.049)
IPA	−0.100***	−0.061*	0.053	−0.284***	−0.201***	−0.145*	−0.108*
	(0.022)	(0.029)	(0.044)	(0.048)	(0.060)	(0.059)	(0.043)
N	475,107	325,442	163,686	372,091	256,166	128,744	372,091

[a] See text for explanation of test A through D. Standard errors in parentheses. Additional regressors are the same as those in Table 4. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Comparison with the lower part of Table 4 shows the estimated incentive effects to be stronger. For instance, the coefficient pertaining to the medium deductible category is −0.137 here, but only −0.042 in Table 4. The likely reason is that the incidence of chronic illness and hence risk selection effects are controlled for to a lesser degree because only one year of data is used for the calculation of the health proxy in test A.

Columns B of Table 6 show the results of pooled estimations over the years 2005 and 2006 only. Here, the health proxy was calculated over two rather than the three previous years prior to averaging. Again, the estimated moral hazard effects tend to exceed the ones of the original model, presumably because the two-year proxy is less effective in controlling for chronic illness than the three-year proxy.

Columns C of Table 6 address the fact that prior HCE is influenced by incentives. As a consequence, a proxy based on prior HCE may make individuals with high deductibles or managed-care type plans appear healthier than they are. To gauge the extent of this potential bias, we recalculated the health proxy by augmenting the observed HCE of individuals in both types of cost-sharing plans, using the estimated coefficients of the incentive effects (lower part of Table 4, full model). As expected, the resulting moral hazard effects are stronger than in Table 4, but not dramatically so.

Finally, a fixed effects specification was used for the second part of the two-part model (column D of Table 4).[11] The dependent variable here is the log of HCE over the years 2004–2006. For the high-deductible plan and the IPA, the estimated incentive effects are close to those in Table 4. For the medium deductible, the direct incentive effect is weaker and insignificant, but the interaction term with the IPA is stronger.

Still another possibility to test the validity of the results presented in Section 6.2 is to track the HCE of individuals who switch between contract types. To an approximation, their personal characteristics are unchanged while contractual incentives are modified.[12] Therefore, it is of interest to compare the difference in HCE prior and after the change, covering the years 2003/04, 2004/05, and 2005/06 with the moral hazard effects displayed in Table 5. For simplicity, only switches away from the baseline (minimum-deductible, FFS) contract are retained (see Table 7). The change in HCE turns out to be symmetrically distributed, permitting estimation of an untransformed linear random effects model. Attribution of annual HCE to contracts hardly causes problems because switches usually take place at the beginning of the year. In order to be able to use several years of data, the health proxy is calculated using the respective previous year only rather than averaging over three years.

Table 7
Analysis of switchers, 2003/04, 2004/05, 2005/06.

Switch from baseline contract	Numbers of switchers	Random effects	
		Coefficient	SE
A. FFS plans			
... to medium DED	2139	−212.99	(130.06)
... to high DED	13,503	−387.99***	(53.27)
B. IPA plans			
... to minimum DED	1905	−260.27	(111.07)
... to medium DED	88	−189.23	(454.35)
... to high DED	805	−219.12	(161.17)

Standard errors in parentheses. Additional regressors are the same as those in Table 4. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

For the switchers from the baseline to a medium-deductible FFS contract, the estimated HCE reduction amounted to CHF 213, which is similar to the CHF 250 of Table 5 (panel A), indicating the attenuation of moral hazard relative to the baseline contract. However, the standard error is large due to the small number of switchers in combination with substantial year-to-year variations in HCE. The switch from the baseline to a high-deductible contract is estimated to be associated with a reduction in HCE amounting to CHF 388, which is again comparable to the figure from Table 5 (CHF 322). This group contains over 13,000 individuals, resulting in a lower standard error and statistical significance.

In the case of supply-side cost sharing, switches from the baseline FFS contract to an IPA option are associated with estimated cost reductions that are again compatible with those evidenced in Table 5. As to the one apparent exception in Table 4 (transition to the medium-deductible, IPA contract), the number of switchers is too low to permit statistical inference.

The evidence compiled in this section comes from two sources. The first consists of several modifications in the estimation of the two-part model. The second source is the analysis of changes in HCE that go along with switching contracts. On the whole, none of the validity tests performed suggests that the effects of supply-side and demand-side cost sharing presented in Table 5 are a mere chance result.

[11] It was not applied to the first part of the two-part model because of the incidental parameters problem (see Section 2).

[12] This statement is only approximately true because the determinants of contract choice must have changed.

6.4. Which types of medical care are most affected by cost sharing?

The encouraging outcome of the validity tests presented in Section 6.3 motivates a more detailed analysis of the effects of demand-side and supply-side cost sharing. The estimation technique described in Section 5 (with the specification in keeping with the full model of Table 4) is applied to general practitioners' services, specialists, drugs, physical therapy, outpatient hospital services, and inpatient hospital services. Therefore, the estimated coefficients reported in Table 8 reflect incentive effects of the two types of cost sharing, with selection effects controlled for (coefficients of the control variables are not shown). For each cost-sharing option, the first line pertains to the first part of the two-part model and second line to the second part, respectively. The interaction terms of deductibles and IPA plans are not shown for brevity. In line with the four-part model advocated by Duan et al. (1982), the probability of observing positive hospital inpatient expenditure is estimated only for the individuals with positive ambulatory care expenditure.

As reported in Table 8, demand-side cost sharing in the guise of a medium deductible significantly reduces the probability of expenditures on GP services, drugs, and hospital outpatient services. For specialized medicine, a significantly positive influence is estimated, which contradicts intuition. As to HCE given that it is positive, there are consistent indications of a reduction effect, which however attains statistical significance in the case of GP services and drugs only.

A high deductible does go along with a decreased probability of all types of care. Moreover, its estimated reduction effect consistently exceeds that of a medium deductible. The impact on hospital inpatient care is surprising as most patients already have HCE in excess of the deductible when entering a hospital. However, patients with high deductibles are less likely to initiate the whole process of diagnostic testing and procedures which ultimately may result in hospitalization (Zweifel, 1992 found similar results using German data). Given positive HCE, there is clear evidence of moral hazard attenuation for GP services, specialist services, and drugs which again exceeds the amount found for the medium deductible.

Turning to supply-side cost sharing, IPA plans exhibit an increased probability of use of GP services, specialist services, and drugs but a decreased probability in the case of physical therapy and hospital services, as expected for the latter. As to the second part of the two-part model, the IPA options are associated with a reduction of expenditure on all types of HCE, with the only exception of hospital outpatient services. Because fees of specialists and hospitals are regulated to be equal for FFS and IPA contracts, these effects are exclusively due to a reduction in quantity.[13]

Summing up, the evidence of Table 8 suggests that supply-side and demand-side cost sharing are effective in different ways. On the demand side, high deductibles primarily seem to lower the likelihood of seeing a GP or a specialist and of using drugs. They also markedly reduce expenditure on GP services, but have no effect on hospital inpatient expenditure. By way of contrast, supply-side cost sharing in the guise of IPA plans is even associated with an increased probability of calling on services of GPs and specialists and of using drugs, presumably due to increased use of preventive care. Its expenditure-reducing effect is concentrated on specialists and inpatient hospital services.

These findings can be compared to the famous RAND Health Insurance Experiment (Manning et al., 1987). The demand-side cost sharing plans in the HIE required the patient to pay a percentage of care out of pocket, up to a stop-loss. In line with the results in Table 8, these plans reduced the probability of incurring any medical expenses, the amount of outpatient expenses, and the probability of inpatient expenses. They did not significantly alter the expenditure for inpatient services.[14] The HMO plan in the HIE was a prepaid group practice. Compared to a FFS plan with no demand-side cost sharing, the HMO plan had a markedly lower hospitalization rate. The effects on hospitalizations are stronger than those measured by this study, which is likely due to stronger financial incentives in the HMO analyzed by RAND, than the IPA analyzed here.

7. Discussion

The aim of this section is to discuss the policy implications of our results, relating them to recent literature. A first salient point is that the estimated absolute cost reduction of CHF 250 (see panel A of Table 5) due to a deductible of CHF 500 rather than 300 exceeds the maximum increase in out-of-pocket expenditure (CHF 200 = 500 − 300). This is confirmed by two other recent studies using Swiss data from earlier years when the minimum deductible was CHF 230 and the next lowest, CHF 400. They both seek to control for risk selection effects. Van Kleef et al. (2008) estimate that raising the deductible from CHF 230 to 400 (i.e. by CHF 170) would serve to reduce expected HCE by CHF 382 (see Table 9). Gardiol et al. (2005) take the maximum deductible of CHF 1500 as their reference point, calculating the incentive effects from there. The transition from the medium deductible of CHF 400 to the minimum of CHF 230 is estimated to generate 'true' savings of CHF 185 (=697 − 512), which again exceeds the out-of-pocket difference of CHF 170. A possible explanation of this 'overshooting' is that patients, who are usually not well informed about the cost of medical care, do not know when they exceed the deductible.

The second point relates to risk adjustment (RA). Note from Table 5 that estimated moral hazard reductions not only fall far short of gross differences in expected HCE as indicated by estimates (1) but are markedly plan-specific. As noted by Van Kleef et al. (2008) and Van Kleef et al. (2006), this varying mix of risk-selection and moral hazard effects poses a great challenge to regulators in a system combining community rating with RA. The issue is the extent to which insurers should be allowed to pass on gross savings to consumers. The appropriate amount seems to be the amount of 'true' savings net of risk-selection effects. Yet, Van Kleef et al. (2006) show that if only very low risks opt for higher deductibles at first, premium reductions reflecting 'true' savings are too small to create incentives for choosing these options.[15] As a remedy, they propose not to entirely net out risk-selection effects for determining allowable premium reductions. Empirical evidence by Van Kleef et al. (2008) reveals that the current RA schemes of the Netherlands and Switzerland do leave room for risk-selection effects in premium

[13] In order to validate our results, we reestimated the second part of the two-part model by OLS on log expenditure. The estimates are close and equal in sign to those in Table 8. The only exception is the coefficient for hospital outpatient services, which is significantly negative in the OLS estimation. It is not a priori clear which estimate is more plausible for this heterogenous patient group. Some patients are chronically ill and in need of repeated procedures (suggesting no effect), while others visit the emergency room for relatively minor ailments (where deductibles might well be effective).

[14] The HIE also included a deductible plan. However, this plan is not directly comparable to the Swiss case because it only applied to outpatient services. It reduced outpatient expenditure and the probability of medical care, but not the probability or amount of inpatient care.

[15] This reflects the Swiss experience after the introduction of voluntary deductibles in 1996.

Please cite this article in press as: Trottmann, M., et al., Supply-side and demand-side cost sharing in deregulated social health insurance: Which is more effective? J. Health Econ. (2011), doi:10.1016/j.jhealeco.2011.10.004

Table 8
Estimation results from the two-part model according to type of care, 2006.

		GP	Specialist	Drugs	Physical therapy	Hospital outpatient	Hospital inpatient
Med. DED	P(HCE > 0)	−0.034***	0.024*	−0.076***	0.004	−0.019*	−0.005
		(0.010)	(0.009)	(0.012)	(0.010)	(0.009)	(0.012)
	HCE|HCE > 0	−0.048***	−0.018	−0.090***	−0.026	−0.045	−0.045
		(0.008)	(0.013)	(0.020)	(0.015)	(0.026)	(0.026)
High DED	P(HCE> 0)	−0.270***	−0.143***	−0.340***	−0.107***	−0.092***	−0.039**
		(0.019)	(0.019)	(0.023)	(0.024)	(0.020)	(0.015)
	HCE|HCE > 0	−0.154***	−0.087***	−0.137***	0.004	0.039	−0.050
		(0.011)	(0.018)	(0.038)	(0.021)	(0.039)	(0.036)
IPA	P(HCE > 0)	0.088***	0.103**	0.087*	−0.049**	−0.058***	−0.066**
		(0.016)	(0.032)	(0.037)	(0.018)	(0.016)	(0.022)
	HCE|HCE > 0	−0.084***	−0.153***	−0.156**	−0.058*	−0.088	−0.116*
		(0.016)	(0.023)	(0.053)	(0.029)	(0.046)	(0.046)
AIC:	P(HCE > 0)	1.042	1.109	0.822	0.820	1.061	0.614
	HCE|HCE > 0	14.41	15.31	15.233	15.254	16.043	19.751
N:	P(HCE > 0)	163,686	163,686	163,686	163,686	163,686	128,744
	HCE|HCE > 0	101,265	86,208	112,941	28,241	46,923	17,205

Standard errors in parentheses. Additional regressors are the same as those in Table 4. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Table 9
Estimated incentive effects of demand-side cost sharing on HCE, in CHF.

Deductible levels	230	400	600	1200	1500
Effects reported by Van Kleef et al. (2008)	−3	−382	−443	−276	−318
Effects reported by Gardiol et al. (2005)	+697	+512	+306	+62	0
Deductible levels	300	500		1000–2500	
Effects reported by this study	0	−250		−322	

CHF 1≈ EUR 0.66

reductions. Their finding is replicated by this study since the maximum allowable premium reductions for high deductibles range between CHF 560 and 1760 (see Table 1), exceeding by far the CHF 322 and 268, respectively that can be attributed to the attenuation of moral hazard (see Table 5).

8. Conclusions

Managed competition in social health insurance aims at creating incentives for insurers to increase efficiency and to respond to consumer preferences while preserving solidarity between high- and low-risk types (Van de Ven et al., 2007). Therefore, it is important to know whether contractual innovations such as deductibles or capitated IPA plans achieve 'true' cost savings rather than merely serving as a means for risk selection. This research measures and compares the impacts of demand-side cost sharing (through voluntary deductibles) and supply-side cost sharing (through prepaid IPA plans) on individual health care expenditure (HCE), controlling for risk-selection effects. The data comes from a large panel of Swiss adults covering the years 2003–2006. Since unobserved health status influences both contract choice and HCE, a proxy is constructed from HCE during the first three years of the observation period, complemented by the residuals from the contract choice equation (the two-stage residual inclusion method proposed by Terza et al. (2008)).

Higher annual deductibles and IPA plans are both found to achieve marked reductions of moral hazard. An increase in the annual deductible by CHF 200 (some EUR 133, from minimum to medium) is estimated to decrease the probability of positive HCE by almost 1 percentage point, while the IPA alternative might even be associated with an increase. In return, it achieves a reduction of positive HCE by some 12 percent, compared to only 7 percent of the medium deductible. Increasing the deductible by CHF 700 (some EUR 466) reduces the probability of reporting HCE by about 3.7 percent and the amount of positive HCE by about 18 percent.

However, this effectiveness of demand-side cost sharing comes at the price of substantial risk-selection effects. Because voluntary cost sharing plans are especially attractive to low risks, such plans might lead to market segmentation and hence higher premiums for high risks. The most favorable ratio of moral hazard attenuation over risk selection is achieved by the IPA with the minimum deductible, amounting to 1:2.76 (CHF 355/980; HCE amounts to CHF 4610 for the baseline contract). The next-best alternative is the medium-deductible FFS contract with a ratio of 1:3.56 (CHF 250/891). According to this criterion, supply-side cost sharing is somewhat more effective than the demand-side alternative.

Still, this research is subject to several limitations. First, since the data set only comprises individuals who were with one and the same insurer from 2003 to 2006, it fails to measure risk-selection effects associated with changes between competing insurers. Second, even 'within' risk-selection effects may not be controlled for perfectly. There is no guarantee that the HCE equation is correctly specified for the three preceding years, a necessary condition for obtaining residuals that serve as good proxies for unobserved health. The same caveat applies to the residuals of the contract choice equations. Thus, estimates of expected HCE reductions achieved by higher deductibles and IPA plans could still be biased. Third, deductible options have price and income effects. As shown by Nyman (1999), only the former should be counted as inefficient consumption. Finally, results relating to IPA plans have limited generality as long as they cannot be linked in detail to the incentives faced by participating health care providers.

Nevertheless, the findings of this study permit one to draw the conclusion that allowing insurers to offer plans with both demand-side and supply-side cost sharing does generate 'true' savings in Swiss social health insurance. After controlling for risk-selection effects, both variants are estimated to achieve marked reductions in moral hazard that can be passed on to consumers in the guise of premium reductions without jeopardizing insurers' solvency.

Appendix A. Estimating a proxy for health status

Table A.10

Table A.10
Estimation of Eqs. (1) and (2).

	$P(HCE_{it} > 0)$ Random effects probit	$HCE_{it} \mid HCE_{it} > 0$ Random effects AR (1)
Greater metropolitan area	0.018	−0.077***
	(0.021)	(0.010)
Affluent community	0.101*	−0.069***
	(0.041)	(0.020)
Regional center	−0.083**	−0.203***
	(0.032)	(0.017)
Rural, mainly industrial	−0.070**	−0.173***
	(0.025)	(0.013)
Rural, agriculture	−0.171***	−0.209***
	(0.028)	(0.015)
Berne city	0.088**	0.037*
	(0.033)	(0.016)
Lucerne city	−0.059*	−0.241***
	(0.028)	(0.014)
Geneva city	0.396***	0.385***
	(0.042)	(0.019)
2004	0.007	0.091***
	(0.012)	(0.005)
2005	0.017	0.089***
	(0.012)	(0.005)
Constant	1.928***	6.911***
	(0.036)	(0.018)
N	253,653	218,208
α	0.0691	0.531
ρ		0.090

Standard errors in parentheses. Additional regressors are age, gender and additional region-specific dummies. α : Fraction of error variance due to l individual-specific term. ρ : Estimated autocorrelation coefficient. $^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Appendix B. Predicting contract choice

Table B.11

Table B.11
Estimation of contract choice in 2006.

	Choice of deductible, ordered probit	Choice of prepaid IPA, probit
Health proxy	−0.341***	−0.219***
	(0.006)	(0.012)
(Health proxy)^2	0.042***	0.055***
	(0.003)	(0.005)
(Health proxy)^3	−0.002***	−0.004***
	(0.000)	(0.001)
Bad credit record	−0.274***	−0.359***
	(0.014)	(0.027)
Years of CSS membership since 1999	−0.042***	−0.052***
	(0.004)	0.007)
Baseline premium	0.018***	0.005***
	(0.002)	(0.001)
Premium reduction for medium DED	0.008	
	(0.010)	
Premium reduction for high DED	0.015***	
	(0.003)	
Premium reduction for IPA		0.013***
		(0.001)
IPA operational in zipcode area		1.436***
		(0.058)
Constant		−3.926***
		(0.287)
Cut points	7.145/7.823	
Log likelihood	−140,618	−27,297
Number of observations	163,686	163,686

Standard errors in parentheses. Additional regressors are age, gender, types of municipalities, region specific dummies, accident coverage, long term care coverage, and youth rebate eligibility. $^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

References

Arrow, K., 1963. Uncertainty and the welfare economics of medical care. American Economic Review 53, 941–973.
Baltagi, B., Wu, P., 1999. Unequally spaced panel data regressions with AR(1) disturbances. Econometric Theory 15, 814–823.
Blough, D., Madden, C., Hornbrook, M., 1999. Modeling risk using generalized linear models. Journal of Health Economics 18, 153–171.
Cameron, A., Trivedi, P., 2005. Microeconometrics, 1st ed. Cambridge University Press, New York.
Cameron, A., Trivedi, P., Milne, F., Piggott, J., 1988. A microeconometric model of the demand for health care and health insurance in Australia. Review of Economic Studies 55, 85–106.
Cardon, J., Hendel, I., 2001. Asymmetric information in health insurance: evidence from the national medical expenditure survey. RAND Journal of Economics 32, 408–427.
Chamberlain, G., 1980. Analysis of covariance with qualitative data. Review of Economic Studies 47, 225–236.
Chiappori, P., Durand, F., Geoffard, P., 1998. Moral hazard and the demand for physician services. European Economic Review 42, 499–511.
Deb, P., Munkin, M., Trivedi, P., 2006. Bayesian analysis of the two-part model with endogeneity: application to health care expenditure. Journal of Applied Econometrics 21, 1081–1099.
Deb, P., Trivedi, P., 2009. Provider networks and primary-care signups: do they restrict the use of medical services? Health Economics 18, 1361–1380.
Dowd, R., Feldman, R., Cassou, S., Finch, M., 1991. Health plan choice and utilization of health care services. The Review of Economics and Statistics 73, 85–93.
Duan, N., Manning, W., Morris, C., Newhouse, J., 1982. A comparison of alternative models for the demand for medical care. In: The Rand Publication Series R-2754-HHS.
Eichner, M., 1998. The demand for medical care: what people pay does matter. AER Papers and Proceedings 88, 117–121.
Ellis, R., McGuire, T., 1993. Supply-side and demand-side cost sharing in health care. Journal of Economic Perspectives 9, 375–396.
Gardiol, L., Geoffard, P., Grandchamp, C., 2005. Separating selection and incentive effects in health insurance. PSE Working Papers 2005-38.
Gardiol, L., Geoffard, P., Grandchamp, C., 2006. Selection and incentive effects: an econometric study of Swiss health-insurance claims data. In: Chiappori, P., Gollier, C. (Eds.), Competitive Failures in Insurance Markets. MIT Press, Cambridge, MA, pp. 81–95.
Gerfin, M., Schellhorn, M., 2006. Nonparametric bounds on the effect of deductibles in health care insurance on doctor visits – Swiss evidence. Health Economics 15, 1011–1020.
Gourieroux, C., Monfort, A., Renault, E., Trogon, A., 1987. Generalized residuals. Journal of Econometrics 34, 5–32.
Hardin, J., Hilbe, J., 2007. Generalized Linear Models and Extensions, 2nd ed. Stata Press, College Station.
Heckman, J.,1976. The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. In: Annals of Economic and Social Measurement. NBER, New York, pp. 120–137.
Jones, A., 2000. Health econometrics. In: Culyer, A., Newhouse, P. (Eds.), Handbook of Health Economics. Elsevier, Amsterdam, pp. 265–344.
Kreier, R., Zweifel, P., 2010. Health insurance in Switzerland: a closer look at a system often offered as a model for the United States. Hofstra Law Review 39, 89–110.
Lamers, L., Van Vliet, R., 2003. Health-based risk adjustment: improving the pharmacy-based cost group model to reduce gaming possibilities. The European Journal of Health Economics 4, 107–114.
Lancaster, T., 2000. The incidental parameter problem since 1948. Journal of Econometrics 95, 391–413.
Lee, L., 1978. Unionism and wage rates: a simultaneous model with qualitative and limited dependent variables. International Economic Review 23, 415–433.
Lehmann, H., Zweifel, P., 2004. Innovation and risk selection in deregulated social health insurance. Journal of Health Economics 23, 997–1012.
Manning, W., 1998. The logged dependent variable, heteroscedasticity, and the retransformation problem. Journal of Health Economics 17, 283–295.
Manning, W., Mullahy, J., 2001. Estimating log models: to transform or not to transform? Journal of Health Economics 20, 461–494.
Manning, W., Newhouse, J., Duan, N., Keeler, E., Leibowitz, A., 1987. Health insurance and the demand for medical care: evidence from a randomized experiment. The American Economic Review 77, 251–277.
Newhouse, J., 2002. Pricing the Priceless: A Health Care Conundrum. MIT Press, Cambridge.
Norton, E., Wang, H., Ai, C., 2004. Computing interaction effects and standard errors in logit and probit models. The Stata Journal 4, 154–167.
Nyman, J., 1999. The economics of moral hazard revisited. Journal of Health Economics 18, 811–824.
OECD, 2006. Switzerland, OECD Reviews of Health Systems.
Schellhorn, M., 2001. The effect of variable health insurance deductibles on the demand for physician visits. Health Economics 10, 441–456.
Terza, J., Basu, A., Rathouz, P., 2008. Two-stage residual inclusion estimation: addressing endogeneity in health econometric modeling. Journal of Health Economics 27, 531–543.
Van de Ven, W., Beck, K., Van de Voorde, C., Wasem, J., Zmora, I., 2007. Risk adjustment and risk selection in Europe: six years later. Health Policy 83, 162–179.

Please cite this article in press as: Trottmann, M., et al., Supply-side and demand-side cost sharing in deregulated social health insurance: Which is more effective? J. Health Econ. (2011), doi:10.1016/j.jhealeco.2011.10.004

Van Kleef, R., Beck, K., Van de Ven, W., Van Vliet, R., 2008. Risk equalization and voluntary deductibles: a complex interaction. Journal of Health Economics 27, 427–443.

Van Kleef, R., Van de Ven, W., Van Vliet, R., 2006. Premium rebate in exchange for a voluntary deductible in social health insurance with risk equalisation: community-rated of risk-rated? Journal of Risk and Insurance 73, 529–550.

Vella, F., 1993. A simple estimator for simultaneous models with censored endogenous regressors. International Economic Review 34, 441–457.

Vella, F., Verbeek, M., 1998. Whose wages do unions raise? A dynamic model of unionism and wage rate determination for young men. Journal of Applied Econometrics 13, 163–183.

Vella, F., Verbeek, M., 1999. Two-step estimation of panel data models with censored endogenous variables and selection bias. Journal of Econometrics 90, 239–263.

Werblow, A., Felder, S., Zweifel, P., 2007. Population ageing and health care expenditures: a school of 'red herrings'? Health Economics 16, 1109–1126.

Winkelmann, R., 2004. Co-payments for prescription drugs and the demand for doctor visits – evidence from a natural experiment. Health Economics 13, 1081–1089.

Wolfe, J., Godderies, J., 1991. Adverse selection, moral hazard, and wealth effects in the Medigap insurance market. Journal of Health Economics 10, 433–459.

Wooldridge, J., 2002. Econometric Analysis of Cross Section and Panel Data. MIT Press, Cambridge, MA.

Zeckhauser, R., 1970. Medical insurance: a case study of the tradeoff between risk spreading and appropriate incentives. Journal of Economic Theory 2, 10–26.

Zweifel, P., 1992. Bonus Options in Health Insurance. Kluwer Academic Publishers, The Netherlands.

Zweifel, P., Manning, W., 2000. Moral hazard. In: Culyer, A., Newhouse, J. (Eds.), Handbook of Health Economics. Elsevier, Amsterdam, pp. 410–459.