What Drives Individual Health Expenditure in Switzerland?[a]

Marcel Bilger and Jean-Paul Chaze[b]

JEL-Classification: C24, D12, I11
Keywords: healthcare demand, health expenditure, hurdle models, Box-Cox

1. Introduction

Analyzing the determinants of healthcare consumption is undoubtedly of primary importance considering the growing concern about health system financing. A comprehensive analysis of this issue would require detailed and reliable individual-level data on healthcare expenditure, along with a wide set of individual characteristics, including health status and health insurance type. Such an ideal data set does not exist in Switzerland, though. The required variables are spread over different sources which will be briefly discussed here.

To begin with, a natural source for healthcare consumption data comes from health insurers, who may provide reliable data on health expenditures based on insured claims. Moreover, the type of expenditure is known by the insurer as well as the insurance coverage. Insurance data has been used by Lehmann and Zweifel (2004) who measured the innovation effect of managed care plans. It is interesting to remark that the authors used past healthcare expenditures in order to infer the missing health status needed in their study to predict healthcare expenditure. In their very close objective of measuring the incentive effect of social health insurance deductibles, Gardiol, Geoffard and Grandchamp (2005) brought a theoretical solution to the missing health status. These authors stated a structural microeconomic model explaining the demand of both health care and health insurance. One of the key elements is that, under certain assumptions,

a This paper derives from the joint IRIS research program of the universities of Geneva and Lausanne and the federal polytechnic school of Lausanne. The program is funded by the CUS and by said institutions.

b Authors's address: Department of Econometrics and Laboratory of Applied Economics, University of Geneva, bd Pont d'Arve 40, 1211 Geneva 4, Switzerland. Email: marcel.bilger@metri.unige.ch and jean-paul.chaze@metri.unige.ch.

the observation of the deductible chosen by the individuals provides information on their health status. Nevertheless, working with health insurance data remains challenging because social health insurers are not allowed to set the premium according to personal characteristics and therefore very little is known about the insured. Typically, only gender, age, and region of residence are recorded. Moreover, insurance data is not designed to be representative of the Swiss population as the risk structure varies between different insurance companies. Finally, even healthcare expenditure is only partially observed, since only bills covered by health insurance are reported, and the insured who consume below their deductible have no necessity to send their claims to the insurance company. Some companies even ask their policyholders not to send their claims until their deductible has been attained, in order to reduce administrative costs.

Another valuable source of data is the Swiss Health Survey (SHS), undertaken by the Swiss Federal Statistical Office (SFSO). This dataset is representative of private households in Switzerland and numerous individual characteristics are recorded, notably including various measurements of health status. Moreover, the type of health insurance plan is also observed, along with healthcare utilization, measured through the frequency of use of various healthcare providers. This survey has been used by Holly, Gardiol, Domenighetti and Bisig (1998) to study the influence of supplemental hospital insurance on inpatient stay. The SHS has also been used extensively to analyze the effect of social health insurance deductibles on the number of doctor visits (see Schellhorn, 2001a, b, 2004; Gerfin and Schellhorn, 2006). It may be noted that moral hazard is not clearly revealed in these papers, whereas studies using health expenditures from insurance data show it more easily (see Werblow, 2002). The main limitation of the SHS comes from the absence of healthcare expenditures, as demand for health care is only measured through frequencies of utilization. Indeed, if we were interested in the actual expenditures, we would have to multiply the number of doctor visits per an average consultation fee. This method is not very accurate and not feasible for other types of expenditures such as inpatient care.

One original feature of the present study is the use of the Swiss Household Income and Expenditure Survey (SHIES) for the analysis of individual healthcare expenditure. This survey, undertaken by the SFSO, is available on a permanent basis since January 2000. Previously, it was only available for 1990 and 1998. Chaze (2005) used a one-off yearly dataset from the 1998 edition to model health expenditure in Switzerland at household level. In the 2000-onwards survey, a sample of approximatively 250 private households is drawn each month. The household members record their revenues and expenses, as well as a number of personal or household characteristics. As regards healthcare consumption, the

SHIES has the advantage of providing detailed information on total expenditure (including co-payments from insurance companies) for different types of healthcare goods or services, that may be covered or not by compulsory health insurance. Moreover, most healthcare expenditures are available at individual level (the remainder being associated to the household). In addition, the SHIES records data on compulsory health insurance deductible and premium, and it also provides means to determine if an individual has an additional health insurance. We will however not make use of these insurance-related factors in this paper, as they require further methodological developments (especially the choice of deductible).

The present study is based on the SHIES from 2000 to 2005, which represent 72 monthly samples. The pooling of these samples provides the opportunity to account for the evolution of prices. In particular, the consumer price index for health represents the evolution of costs for health care. Thus, the individuals' healthcare consumption will be measured through real health expenditure, using that price index as a deflator. However, it is important to note that the relative price for health care may only account for part of the "price effect". In fact, the price for health care truly payed by the individuals is endogenous, as it depends on consumption through insurance coverage and the impact of deductible and co-payment. The monthly data available in the SHIES does not allow to compute out-of-pocket health expenditure, as reimbursement from insurance will often enter the household's budget several months after payment. The relative price of healthcare thus only represents the exogenous component that enters the formation of the individuals' prices.

The opportunity to use external information is further enhanced by geographic dispersion. In particular, the *canton* of residence, which plays an important part in the Swiss health system, may be used to account for supply-side factors, such as the densities of physicians or other providers.

Some limitations of the dataset may however be stressed. At first, since expenditure data are collected over a period of one month only, it is consequently prone to infrequent consumption bias, especially with respect to hospital care. Moreover, the SHIES does not contain any measurement of health status, which certainly is the most important factor of healthcare consumption. Therefore, the study aims to relate healthcare consumption to the factors that explain consumer behavior given health status, as well as those that influence health status itself. Although we will not attempt to isolate each aspect (factors such as age, gender or income are likely to influence both of them), this reduced form approach remains relevant from a policy point of view.

The econometric model estimated in this study is a consumption function for healthcare, homogeneous of degree 0 with respect to prices and income, that relates real expenditure to factors such as income, price, age, gender, etc. The methodology used is the Box-Cox indirect censoring model (see CHAZE, 2005) which accounts for the high number of null expenditures typically found in such data as well as the strong asymmetry of positive healthcare expenditures. This modeling approach is presented in section 2. Section 3 presents the dataset and variable definitions, while section 4 shows the estimation of the individual healthcare consumption model. Section 5 concludes.

2. Methodology

Individual or household data on health expenditure generally involves two difficulties that prevent the use of traditional regression methods. Unless the time frame is very long or the population is selected (i.e. patients), the first issue raised is the presence of a large amount of zero expenditures, that calls for a limited dependent variable approach. The second issue comes from the heavily skewed distributions usually exhibited by data on health-related consumption or expenditure, which rule out the normality assumption. The literature on health econometrics (see JONES, 2000, for a review), generally considers two approaches. The most commonly used in applied research relates to two-part models, which are based on the following decomposition (valid as long as each component is defined):

$$E(y \mid x) = \Pr(y > 0 \mid x) E(y \mid y > 0, x). \tag{1}$$

These models rely on the assumption that both parts on the right hand side, i.e. the probability of consuming and the consumption level of consumers, depend on different sets of parameters, thus allowing separate estimation. One advantage of two-part models is that keeping each part in a single dimension framework allows full flexibility in the treatment of skewed outcomes, by using for instance GLM (see BLOUGH, MADDEN and HORNBROOK, 1999) or generalized gamma distributions (see MANNING, BASU and MULLAHY, 2005). However, although the separability assumption does not imply independence of the selection and conditional consumption processes (as pointed out by DUAN, MANNING, MORRIS and NEWHOUSE, 1984), it may prove overly restrictive in some situations. The alternative approach is the selection or hurdle approach, which involves a full specification of the link between the aforementioned processes.

In that context, Chaze (2005) developed a family of Box-Cox censoring models for limited dependent variables, based on the two-parameter Box-Cox transformation (Box and Cox, 1964):

$$T(y)=T(y;\lambda,\gamma)=\begin{cases}\dfrac{(y+\gamma)^{\lambda}-1}{\lambda} & \text{if } \lambda\neq 0\\ \log(y+\gamma) & \text{if } \lambda=0\end{cases}. \tag{2}$$

The shape parameter λ accounts for the degree of skewness (bringing the transformed variable close to normality), while the location parameter γ allows linear translations of the data. The standard Box-Cox transformation is obtained by setting $\gamma=0$.

The interest of the location parameter for modeling purposes is that it allows generating zero responses as corner solutions from the consumption process. Chaze proposes three complementary specifications for Box-Cox censoring models. The first is a single-equation model generalizing the standard Tobit, which is essentially a reformulation of the Box-Cox standard Tobit by Lankford and Wyckoff (1991). The second and third specifications involve a selection equation in addition to the consumption equation. For the Box-Cox indirect censoring model, the location parameter is set to zero, so that all the zero consumptions are generated by the selection equation, while in the Box-Cox double censoring model, zero consumptions are allowed to be generated by both equations.

This family of models provides more flexibility than previous specifications based on the standard Box-Cox transformation, notably Box-Cox double hurdle models by Yen (1993), Yen, Dellenbarger and Schupp (1995), and Jones and Yen (2000). More generally, these models represent a good trade-off between the ability to address the nonnormality issues in the consumption part while keeping the link with the selection process and allowing the possibility to test its significance. In Chaze (2005), the Box-Cox double censoring model is applied to household health expenditure in Switzerland. The underlying data comes from a special questionnaire only available in the 1998 edition of the SHIES, where participating households were instructed to report their healthcare expenditures of CHF 150 or more from the previous 12 months. The presence of a data censoring threshold and heavy skewness of the data required full use of the flexibility of the Box-Cox double censoring model. With the monthly expenditure data used here, where households are instructed to report all their expenditures, there is no a priori necessity for the model to be able to generate zeros as corner solutions from the consumption equation. Therefore we restrict ourselves to the Box-Cox indirect censoring model, which sets the location parameter $\gamma=0$.

The Box-Cox indirect censoring model is defined by

$$\begin{cases} y_1^* & = & x'\beta_1 + \sigma u_1 \\ y_2^* & = & x'\beta_2 + u_2 \end{cases}, \tag{3}$$

and

$$y = \begin{cases} T^{-1}(y_1^*) & \text{if } y_2^* > 0 \\ 0 & \text{otherwise} \end{cases}, \tag{4}$$

with y being the observed consumption, y_1^* a latent variable underlying the amount of positive consumption in the transformed scale, y_2^* a latent variable underlying the selection process, x a vector of independent variables, β_1 and β_2 the parameter vectors associated to x in each equation, and $\sigma > 0$ the standard error parameter. The errors terms u_1 and u_2 are assumed to follow a standard bivariate normal with correlation parameter ρ, and the distribution of u_1 is truncated between B_1 and B_2, where:

$$B_1 = B_1(x;\beta_1,\lambda,\sigma) = \begin{cases} -(x'\beta_1 + 1/\lambda)/\sigma & \text{if } \lambda > 0 \\ -\infty & \text{if } \lambda \leq 0 \end{cases}, \tag{5}$$

and

$$B_2 = B_2(x;\beta_1,\lambda,\sigma) = \begin{cases} +\infty & \text{if } \lambda \geq 0 \\ -(x'\beta_1 + 1/\lambda)/\sigma & \text{if } \lambda < 0 \end{cases}. \tag{6}$$

These bounds insure that $T^{-1}(y_1^*)$ is defined, whatever λ.

The parameters of the model, β_1, β_2, λ, ρ and σ, are estimated by maximum likelihood (see CHAZE, 2005, for the expression of the log-likelihood).

As we are not interested to infer on consumption in the transformed scale, we need to compute formulas for expected consumption $E(y \mid x)$, probability of consuming $Pr(y > 0 \mid x)$, and expected consumption when consuming $E(y \mid y > 0,x)$, conditionally on the values of the independent variables x, and derive the corresponding marginal effects. We have (provided the integral converges):

$$\begin{aligned} E(y \mid x) &= Pr(y > 0 \mid x)E(y \mid y > 0, x) \\ &= \int_0^{+\infty} yPr(y > 0 \mid x) f(y \mid y > 0, x)dy \end{aligned} \tag{7}$$

Once the "unconditional" predictor $\hat{E}(y \mid x)$ is computed (replacing the unknown parameters by their estimated values), the estimated probability $Pr(y > 0 \mid x)$ may be used to compute the conditional predictor:

$$\hat{E}(y \mid y > 0, x) = \hat{E}(y \mid x) / \hat{P}r(y > 0 \mid x). \tag{8}$$

We hereby present formulas for unconditional expectations $E(y \mid x)$, that need to be computed numerically. Note that we choose the latent variable y_1^* as the integration variable instead of the error term u_1, in order to avoid the bounds of the integral depending on x. Consider the auxiliary function:

$$\psi(u_1, u_2, \rho) = \varphi(u_1)\Phi\left([u_2 + \rho u_1] / \sqrt{1-\rho^2}\right). \tag{9}$$

We have, for $\lambda > 0$:

$$\begin{aligned} E(y \mid x) = & \frac{1}{\sigma\Phi\left((x'\beta_1 + 1/\lambda)/\sigma\right)} \\ & \times\left(\int_{-1/\lambda}^{+\infty} (\lambda y_1^* + 1)^{1/\lambda} \psi([y_1^* - x'\beta_1]/\sigma, x'\beta_2, \rho) dy_1^*\right), \end{aligned} \tag{10}$$

for $\lambda < 0$:

$$\begin{aligned} E(y \mid x) = & \frac{1}{\sigma\Phi\left((x'\beta_1 - T(y_s))/\sigma\right)} \\ & \times\left(\int_{-\infty}^{T(y_s)} (\lambda y_1^* + 1)^{1/\lambda} \psi([y_1^* - x'\beta_1]/\sigma, x'\beta_2, \rho) dy_1^*\right), \end{aligned} \tag{11}$$

(note that in the $\lambda < 0$ case, we need to set an upper bound y_s on consumption, in order to insure the convergence of the integral), and for $\lambda = 0$:

$$E(y \mid x) = exp(x'\beta_1 + \frac{1}{2}\sigma^2)\Phi(x'\beta_2 + \rho\sigma). \tag{12}$$

On the other hand, the probability of consuming is simply given by:

$$Pr(y > 0 \mid x) = \Phi(x'\beta_2). \tag{13}$$

Formulas for marginal effects are obtained by analytical derivation. They also need to be computed numerically. Standard errors for the predicted values and marginal effects may be obtained by the delta method.

3. Data

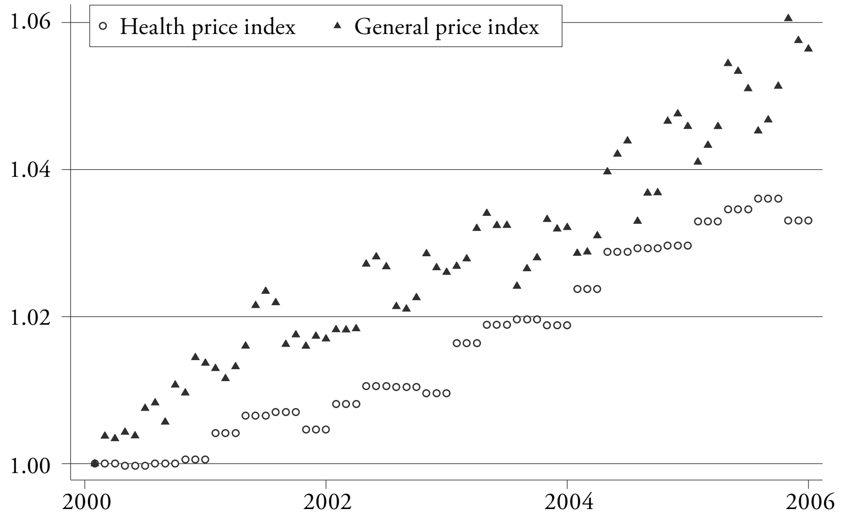
As mentioned in the introduction, the micro data analyzed here comes from the SHIES, which is made of monthly cross-sectional samples and has been carried out by the SFSO on a regular basis since January 2000. In the present study, 72 monthly samples collected over the period 2000-2005 have been combined, giving rise to one global sample of 20’940 households. In order to obtain an independent sample of adult individuals, one person, aged 18 or more, has been randomly drawn from each household, and a few observations have been dropped because of missing values on key variables. The resulting sample amounts to 20’925 individuals aged 18 or more. The children’s consumption, which follows specific patterns (especially at very young age), is not considered in this study.

Our measurement of healthcare consumption at a given period is computed as the sum of all expenditures on healthcare over a period of one month. It is important to point out that SHIES data relates to total cost of healthcare and not out-of-pocket expenditure. In most cases, health expenditure is observed at individual level, as only 10.4% of the expenditures were reported at household level. Our approach is to allocate the latter to household members proportionally to their observed health expenditure. To account for the evolution of prices, total healthcare expenditure has been divided by the consumer price index for healthcare provided by the SFSO whose reference period has been set to January 2000. The reference for all monetary values (in CHF) given in this study will thus be January 2000. Figure 1 shows the evolution of both health and general price indexes over the period of interest. One can see that the price of healthcare has increased more slowly than the general price index. One may also remark that the price index for health is updated every three months only.

Table 1 provides descriptive statistics for individual healthcare consumption, as well as for the quantitative explanatory variables used in the model. In the first line of the table, one can see that the sample mean of individual healthcare consumption amounts to CHF 188.1. The very small median (CHF 8.4) reflects the high percentage of non-consumers during the month of survey (43.3%). The second line of the table presents the same statistics for positive health consumption only. This conditional variable has a sample mean of CHF 331.8 and median of CHF 75.4. Moreover, the conditional standard deviation is very large and the quartiles reveal that the variable distribution is highly skewed.

The income variable has been computed as the total household expenditure for consumer goods and services, minus purchases of motor vehicles, and minus healthcare expenditures, in order to avoid simultaneity bias caused by the impact

Figure 1: Monthly Health and General Price Indexes



of high infrequent health related expenses on the monthly household budget. This variable acts as a proxy for household permanent income, and leads to better results than variables based on monthly household income. The income variable is then divided by the general consumer price index (reference January 2000) and by an equivalence scale to account for household composition.[1] With this definition, the average real income amounts to CHF 5105.8. The sample mean for age equals 47.3 years, with a standard deviation of 16 years. It may be noted that the SHIES is not representative of the whole population because it only includes individuals living in private households. Therefore the elderly living in a nursing home are not observed, even though their healthcare consumption is typically higher than those living in a private home. The last three variables in Table 1 relate to the number of physicians, hospital beds and pharmacies per 100'000 inhabitants in the *canton* of residence. These important figures of the health system supply are provided by SFSO on a yearly basis.

1 The equivalence scale factor is computed as the square root of the sum of the number of adults and half the number of children.

Table 1: Statistics on Quantitative Variables

	mean	s.d.	Q1	Q2	Q3
Health consumption	188.1	871.9	0.0	8.4	103.1
Positive health consumption	331.8	1137.3	19.4	75.4	279.4
Income	5105.8	2827.7	3430.4	4564.4	6078.9
Age	47.3	16.0	35.0	45.0	59.0
Physician density	197.1	50.9	155.2	198.8	225.1
Hospital bed density	604.7	159.2	532.2	606.9	650.5
Pharmacy density	23.6	13.5	14.5	18.3	36.3

Table 2 presents the qualitative explanatory variables along with summary statistics on the distribution of healthcare consumption for each category. Column 2 gives the frequency of each category. The next two columns show the mean and the percentage of positive healthcare consumptions respectively. The last four columns refer to the distribution of positive healthcare consumption, the conditional mean and quartiles being reported. Household type is the first qualitative variable considered. It can be seen that lone persons and couples without children have the highest healthcare expenditure, which is mainly due to a higher conditional mean. As for category "other", it refers to households with individuals other than a parent or a child. The next qualitative variable refers to the degree of urbanization of the place of residence. Individuals living in the countryside turn out to spend less on healthcare than those from other categories. Another qualitative variable is the linguistic region, where Italian and Romansh speaking regions have been grouped because of the very small frequency of the latter. Next, a binary variable indicates if the individual has Swiss of foreign nationality. Swiss individuals appear to spend more on healthcare than foreigners do. The level of education (in three categories) appears to have a positive impact on healthcare consumption. Finally, socioprofessional status has been divided in eight categories. Most pensioners are retired individuals and therefore their healthcare consumption is high. Table 3 presents analogous information as table 2 for quantitative variables divided in classes. It can be seen that income has a strong positive impact on healthcare consumption. Demographic variables show that men have a lower healthcare consumption than women, which is predominantly due to a lower frequency of consumption. Finally, all density variables tend to increase consumption, mainly due to a larger conditional mean of real healthcare expenditure.

Table 2: Qualitative Explanatory Variables and Healthcare Expenditure

				Positive health expenditure			
	Freq.	Mean	%Pos.	Mean	Q1	Q2	Q3
Household type							
Lone person	25.9	275.2	65.7	419.0	25.9	98.4	348.0
1 parent, 1 child	2.3	160.8	59.7	269.4	19.5	71.8	199.4
1 parent, ≥ 2 children	2.1	129.2	54.0	239.3	12.3	50.2	252.2
Couple without child	31.4	227.2	61.3	370.8	22.1	90.4	314.9
Couple, 1 child	11.1	117.2	47.8	245.2	15.0	52.6	216.2
Couple, 2 children	17.1	91.4	47.1	194.0	13.6	44.9	168.4
Couple, ≥ 3 children	7.1	99.0	45.9	215.9	10.5	41.8	200.9
Other	2.9	108.3	42.6	254.4	14.9	53.5	256.8
Urbanization							
Countryside	25.7	151.9	53.2	285.3	16.5	60.0	230.5
Small town	0.9	219.6	62.9	349.1	25.1	79.5	294.9
Suburb	46.8	198.2	57.1	347.2	19.7	77.6	285.7
City center	26.6	204.1	59.1	345.4	21.8	85.8	305.8
Linguistic region							
German	69.7	189.8	57.3	331.1	18.1	69.5	269.8
French	21.4	193.6	56.3	343.8	21.6	83.8	282.4
Italian or Romansh	8.8	161.1	52.6	306.2	24.4	106.2	338.1
Nationality							
Swiss	88.8	194.0	58.0	334.6	19.6	75.5	276.7
Foreign	11.2	141.3	46.5	303.8	16.4	74.4	298.6
Education							
Secondary I or less	14.0	177.8	52.1	341.1	21.1	84.1	321.7
Secondary II	71.4	188.6	57.3	329.0	18.7	69.9	264.7
Tertiary	14.6	195.5	58.0	337.2	21.5	94.6	306.9
Socioprofessional status							
Employee	59.0	139.7	53.3	261.8	17.3	60.6	220.1
Self-employed	7.1	191.5	53.5	358.0	16.4	59.0	220.3
Farmer	1.1	58.7	42.4	138.5	10.0	23.5	122.6
Unemployed	1.3	144.5	47.8	302.4	15.0	46.3	167.0
Student	1.9	144.3	48.5	297.7	12.0	38.7	190.4
Homemaker	8.6	151.2	59.5	254.1	17.2	63.1	251.7
Pensioner or other	21.1	350.8	68.0	516.1	35.7	138.9	464.4

Table 3: Quantitative Exploratory Variables and Healthcare Expenditure

				Positive health expenditure			
	Freq.	Mean	%Pos.	Mean	Q1	Q2	Q3
Income class							
< 3000	16.5	179.3	52.7	340.4	18.1	72.5	303.2
3000–4000	21.1	158.5	53.4	296.6	16.5	61.4	240.9
4000–5000	21.0	192.2	55.2	348.1	17.5	70.7	263.7
5000–6000	15.4	168.7	57.4	294.1	20.0	70.7	252.7
6000–8000	15.9	199.0	62.3	319.4	24.1	87.1	279.9
> 8000	10.1	268.0	63.2	424.0	25.3	111.0	357.4
Age class for women							
18–29	7.4	137.7	57.0	241.6	16.9	54.6	186.5
30–39	13.3	152.3	59.4	256.4	17.3	59.3	212.2
40–49	10.4	160.2	62.3	257.0	18.7	64.8	240.2
50–59	8.9	263.8	67.2	392.4	24.7	95.3	318.5
60–69	7.6	329.3	71.7	459.2	31.7	124.1	428.2
≥ 70	6.4	370.9	72.8	509.6	40.2	144.4	441.9
Age class for men							
18–29	5.5	66.0	37.2	177.2	10.6	24.8	125.5
30–39	11.4	87.9	41.6	211.3	12.6	43.5	194.4
40–49	10.7	130.8	44.2	295.7	13.5	47.8	187.3
50–59	7.7	164.3	50.3	326.8	18.8	85.0	261.8
60–69	6.1	237.4	59.3	400.3	21.0	99.9	383.0
≥ 70	4.7	301.2	64.6	466.0	35.4	145.5	461.8
Physician density class							
< 150	16.9	164.1	55.4	296.2	18.6	75.0	268.1
150–200	35.6	183.7	55.6	330.3	18.3	68.0	263.8
200–250	41.2	195.3	57.8	338.2	19.9	79.3	279.9
> 250	6.3	229.7	59.1	388.3	22.2	91.3	368.1
Hospital density class							
< 500	17.5	165.9	54.7	303.6	16.7	58.5	223.5
500–600	29.3	183.4	56.9	322.6	17.9	66.6	240.9
600–700	39.6	194.9	57.5	338.9	20.8	84.8	294.0
> 700	13.6	206.8	56.5	365.9	21.8	92.9	358.5
Pharmacy density class							
< 15	26.9	183.4	55.4	331.3	17.4	63.8	257.7
15–30	43.1	188.1	58.4	322.3	18.3	71.0	269.9
> 30	30.0	192.2	55.5	346.6	22.6	91.6	304.2

4. Results

In this section, we apply the Box-Cox indirect censoring model (see section 2) to healthcare consumption of adult individuals living in private households in Switzerland, using the 2000–2005 sample described in the previous section. Explanatory variable selection has been performed by placing all variables in both the consumption and selection equations. It may be noticed that due to the nonlinearity of the system, exclusion restrictions are not required. The selection process removes one variable in one equation (variables that generate more than one parameter in each equation being considered as a whole) at each step, until all remaining variables are significant at the 5% level, and the best model is selected according to Akaike's information criterion (AIC). It is worth mentioning that some alternative explanatory variables have also been tested. This is notably the case for factors that may be introduced at individual or household level, using the reference person's characteristics instead of the individual's in the latter case. The variables that we considered with this respect are socioprofessional status and education level, while the individual's marital status has been tested as a substitute for household type. Complete variable selection has been performed for all possible combinations of these alternative variables, with the best overall model again selected according to AIC. Finally, it should be mentioned that the variables which are theoretically important have been kept in the model, irrespective of their statistical significance. These are the logarithm of relative price for healthcare, the logarithm of equivalent income, gender, a degree 5 polynomial of age, with interaction terms for the latter two variables.

The result of alternative explanatory variable selection is that socioprofessional status and education level are more influential at individual level (probably due to the fact that children are not considered in the study), whereas household type is more relevant than individual marital status. As to other explanatory variables, it appears that neither the linguistic region nor the degree of urbanization are significant in any of the two equations. It is likely that the effect of these two geographic variables on healthcare consumption has been captured by health supply factors. Moreover, dummy variables indicating the month of survey have also been introduced into the model. It turns out that the month of survey appears significant in neither the consumption nor the selection equation. That seems surprising at first glance, considering seasonal diseases and the fact that the insured may have the incentive to consume more at the end of the year when their health insurance deductible has been reached. However, this result is probably explained by time delays in billing. Indeed, the time elapsed between actual consumption and payment is highly variable, and is likely to mask any seasonality effect.

Table 4 shows the results of the estimation of the individual health consumption model. The first two columns of the table give the estimated coefficients of the expenditure equation along with their standard errors, and the last two columns give the same information for the selection equation. Let us comment on nonlinear parameters first. One can see that the shape parameter λ is slightly positive and significantly different from zero at the 5% level. This indicates that the logarithmic transformation would induce a distribution that is slightly skewed to the left and that the use of the Box-Cox, while remaining fairly close to the logarithm, improves the fit of the model. The correlation parameter ρ appeared not to be significantly different from 0, and has thus been constrained. In that case, the Box-Cox indirect censoring model reduces to a particular case of a two-part model, each equation relying on different parameters.

A first result concerning explanatory variables is that the relative price for healthcare appears significant in neither the consumption nor selection equation. This result may be attributed to the fact that individuals are not very sensitive to the actual price of healthcare, as they pay only part of the cost thanks to their insurance coverage. Income has a highly significant positive impact in both equations. With all the variables set at their sample mean, income elasticity of unconditional healthcare consumption amounts to 0.70, which breaks up into 0.24 for participation and 0.47 for conditional consumption. As remarked by Chaze (2005), even though a higher income likely contributes to reducing the needs for healthcare, an increase in prevention and quality of care may explain this effect of income. The results found in the present study bring further support to this interpretation. Indeed, income elasticity is lower for participation, which may reflect the fact that an increase in income will draw some healthy individuals to recourse to preventive healthcare, while some sick persons will gain access to healthcare. On the other hand, income elasticity of conditional consumption is higher, reflecting the demand for quality care when income increases.

As to the impact of the age polynomial, it appeared that although the fourth and fifth powers are not jointly significant at the 5% level in the selection equation, they presence improves the fit of the model according to AIC, and are therefore kept in both equations. Figures 2, 3, and 4 show, for each gender, the evolution according to age of expected unconditional consumption, probability of consuming, and expected conditional consumption, with all the other variables set at their sample mean. Interestingly, there is a marked difference between men and women. Women exhibit higher consumptions for all ages, which is due to both higher participation and conditional consumption, except in the latter case for older ages. Moreover, women's health consumption attains a local maximum at the age of 32, obviously associated to maternity, then slightly decreases

Table 4: Estimation of the Individual Health Expenditure Model

	Expenditure equation		Selection equation	
	Estimate	s.e.	Estimate	s.e.
Log. of relative health price	−0.63088	3.45280	−0.15624	1.84072
Log. of equivalent income	0.49841	0.04363	0.34216	0.02338
Age	0.78190	0.33633	0.34802	0.18345
Age^2	−0.03369	0.01447	−0.01383	0.00801
Age^3	6.95E−04	2.95E−04	2.60E−04	1.66E−04
Age^4	−6.82E−06	2.88E−06	−2.30E−06	1.63E−06
Age^5	2.57E−08	1.07E−08	7.78E−09	6.15E−09
Male	8.47920	4.34971	3.89781	2.17587
Male × *Age*	−1.10664	0.49654	−0.49941	0.25287
Male × *Age*^{2}	0.05078	0.02137	0.02105	0.01106
Male × *Age*^{3}	−0.00109	4.36E−04	−4.17E−04	2.29E−04
Male × *Age*^{4}	1.10E−05	4.24E−06	3.93E−06	2.26E−06
Male × *Age*^{5}	−4.23E−08	1.58E−08	−1.42E−08	8.54E−09
Household type				
1 parent, 1 child	−0.17818	0.11319	−0.08871	0.06218
1 parent, ≥ 2 children	−0.30778	0.12653	−0.23788	0.06554
Couple without children	−0.25570	0.04432	−0.15875	0.02515
Couple, 1 child	−0.35448	0.06627	−0.33836	0.03354
Couple, 2 children	−0.49825	0.06072	−0.33215	0.03077
Couple, ≥ 3 children	−0.51186	0.08229	−0.33847	0.04041
Other	−0.46531	0.11865	−0.50830	0.05582
Education				
Secondary II	−0.06274	0.05410	0.15496	0.02826
Tertiary	0.05729	0.07149	0.18972	0.03755
Socioprofessional status				
Selfemployed	−0.05455	0.06948	0.00249	0.03582
Farmer	−0.20083	0.18853	0.02762	0.08680
Unemployed	−0.24556	0.16256	−0.07418	0.07957
Student	0.05610	0.14979	0.13976	0.07300
Homemaker	0.07585	0.06532	0.05495	0.03524
Pensioner or other	0.46894	0.06902	0.16250	0.03864
Foreigner	0	constr.	−0.14642	0.02907
Pharmacy density	0.00780	0.00128	0	constr.
Constant	−6.88821	2.99733	−5.87359	1.60970
λ	0.00891	0.00440		
ρ	0	constr.		
σ	1.81839	0.03626		

Figure 2: Expected Healthcare Consumption According to Age

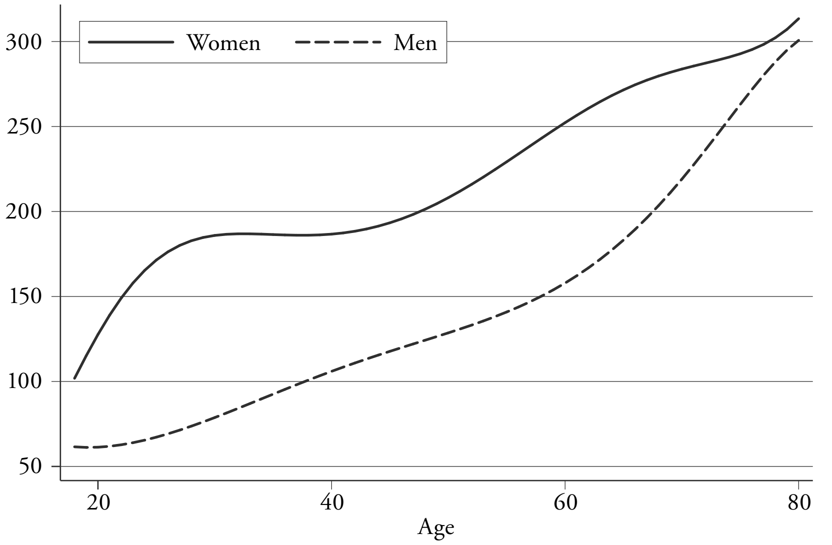


Figure 3: Expected Probability of Consuming According to Age

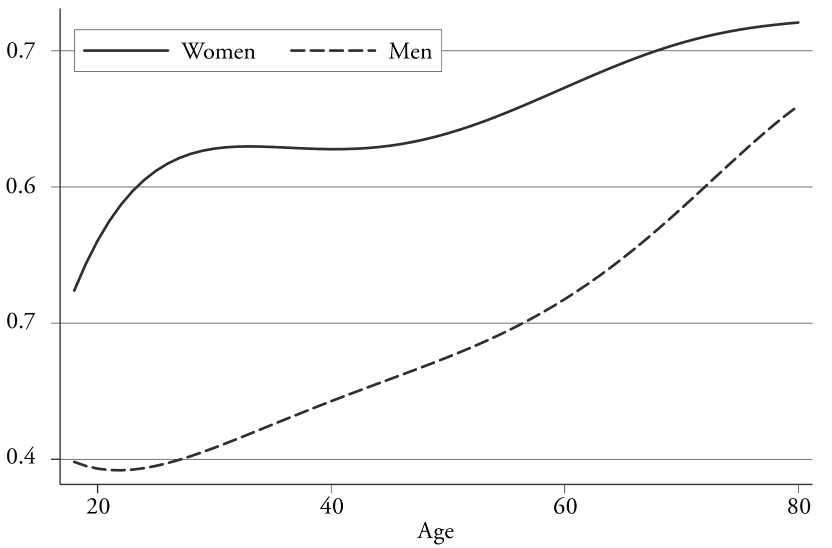
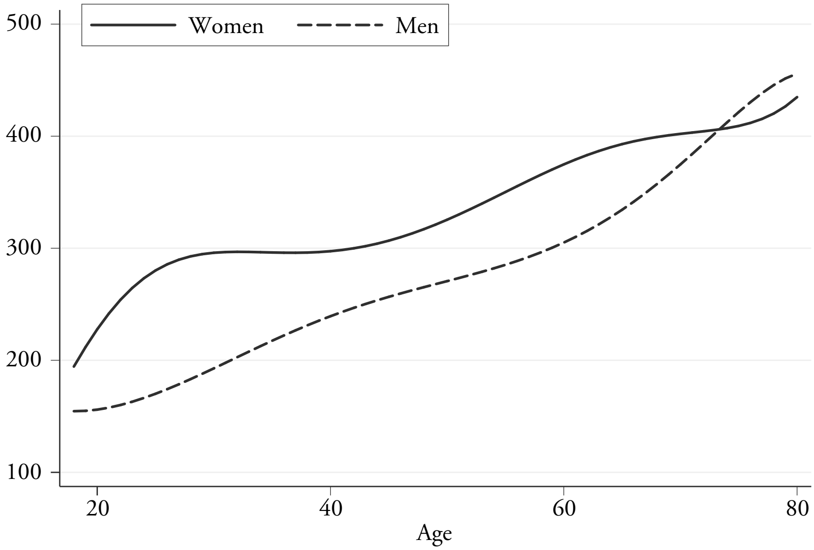


Figure 4: Expected Conditional Healthcare Consumption According to Age



until the age of 38, until it starts increasing again due to aging. As to men, their healthcare consumption according to age exhibits a simpler pattern. Indeed, conditional healthcare consumption is steadily increasing. It may be noted though, that participation starts by slightly decreasing until the age of 22, and then never stops rising. Moreover, the increase rates for both participation and conditional health consumption accelerate around the age of 60, which makes men's healthcare consumption grow faster than women's at older ages.

Table 5 gives the marginal effect of the explanatory variables on expected healthcare consumption, probability of consumption, and expected conditional consuming respectively. The marginal effects presented have been computed at the sample mean of all explanatory variables. For qualitative variables, in addition to marginal effects, one may also compute variations of the expected values obtained by replacing one category by another, with the remaining variables set at their sample mean. Such information will sometimes be given to illustrate specific aspects.

Household type has a surprisingly strong impact on healthcare consumption for adults. Indeed, individuals living alone have both a higher probability of consuming and a higher expected conditional consumption than those living in larger households. Moreover, the number of children tends to reduce healthcare

Table 5: Marginal Effects of the Explanatory Variables

	$E(Y\|x)$		$P(Y>0\|x)$		$E(Y\|Y>0,x)$	
	m.e.	s.e.	m.e.	s.e.	m.e.	s.e.
Household type						
1 parent, 1 child	–38.218	19.209	–0.035	0.024	–49.045	31.186
1 parent, ≥ 2 child.	–75.764	21.430	–0.093	0.026	–84.718	34.978
Couple, no child.	–58.468	7.769	–0.062	0.010	–70.383	12.405
Couple, 1 child	–94.676	11.478	–0.133	0.013	–97.574	18.499
Couple, 2 child.	–116.555	10.830	–0.130	0.012	–137.146	17.297
Couple, ≥ 3 child.	–119.422	14.333	–0.133	0.016	–140.894	23.230
Other	–131.663	20.085	–0.200	0.022	–128.081	32.861
Education						
Secondary II	7.985	9.101	0.061	0.011	–17.268	14.918
Tertiary	30.850	12.062	0.074	0.015	15.769	19.681
Socioprofessional status						
Selfemployed	–8.286	11.678	0.001	0.014	–15.016	19.135
Farmer	–28.382	31.284	0.011	0.034	–55.278	51.924
Unemployed	–47.136	27.183	–0.029	0.031	–67.593	44.797
Student	24.912	25.000	0.055	0.029	15.443	41.227
Homemaker	18.248	11.050	0.022	0.014	20.877	17.986
Pensioner or other	92.412	12.002	0.064	0.015	129.079	19.274
Foreigner	–16.861	3.386	–0.057	0.011	0	0.361
Pharmacy density	1.227	0.206	0	0	2.148	0.361

consumption of the parents. For instance, there is a decrease of CHF 70 in unconditional expected healthcare expenditure between a single person and a couple without children, and a further decrease of CHF 34 between a couple without children and a couple with one child. One possible explanation of these results is that couples, and to a larger extent families, have a healthier way of life and take more care of each other, or that parents cut in their own health consumption to provide more care to their children. A simultaneity bias may also be induced by the fact that persons that suffer long diseases or disabilities may have a higher probability of living alone.

Education is also shown to influence healthcare consumption. In fact, a level of education higher than secondary I tends to increase participation. On the other hand, only the difference between secondary II and tertiary education has a significant influence on conditional healthcare consumption. As to socioprofessional status, the pensioners happen to have the highest healthcare consumption, first through higher participation but mainly through higher conditional consumption. For instance, there is an increase of CHF 108 for the expected healthcare consumption between a wage earner and a pensioner. Note however that a simultaneity bias may be present due to pensioners of the disability insurance. Next, it appears that foreigners have a lesser participation than the Swiss. Nevertheless, nationality does not play any part in the determination of conditional healthcare consumption.

Finally, it has been found that all three healthcare supply variables do not influence participation, but increase conditional healthcare consumption. Moreover, the three supply variables are strongly correlated one to another, and when they are all put together into the model, only pharmacy density appears to be significant at the 5% level. This factor also leads to the best AIC when only one of the densities is included. This statistical relationship between healthcare supply and consumption is consistent with the theory of induced demand. According to it, asymmetric information allows healthcare providers to increase the quantity of care given to each patient, in order to maintain their income level when the number of providers raises and the number of patients remains constant. It is somehow surprising that pharmacy density appears more relevant than physician density to explain conditional healthcare consumption, considering that physicians influence all types of healthcare consumption, and might have the incentive to prescribe more expensive drugs, medical tests, and surgery.

5. Conclusion

In this paper, SHIES data has been used to analyze individual healthcare consumption in Switzerland by means of the family of selection/hurdle models developed by Chaze (2005). It is shown that SHIES data makes it possible to reveal numerous factors of individual healthcare consumption. First of all, age and gender are shown to have a strong impact on healthcare consumption. Women's consumption is greater than men's, especially during maternity years. Apart from that, consumption tends to increase steadily with age. Several household characteristics also prove to be relevant in explaining individual healthcare consumption. Household structure is such a factor, as it appears that lone individuals tend to consume more healthcare than those living in larger households, and also that the consumption of adults tends to decrease with the number of children they have. The other variable measured at household level is permanent income. Health care consumption is shown to have an income elasticity of 0.70 (0.24 for participation and 0.47 for conditional consumption). However, one should keep in mind that household income has to be corrected by an equivalence scale factor. The choice of which may influence the results of the model, especially income elasticities and household structure coefficients. In addition, education level, nationality, and socioprofessional status of the individual also influence healthcare consumption.

It is also interesting to note that the estimated model is consistent with the theory of induced demand. Indeed, all three healthcare supply variables tested, namely the densities of hospital beds, physicians and pharmacies, are shown to increase conditional healthcare expenditure but not participation. Nonetheless, we should be very cautious before concluding to the presence of induced demand, because degree of urbanization is strongly correlated with healthcare supply variables, and drops when the latter are included in the model. Therefore health supply variables will also measure the effect of urbanization.

On the other hand, some variables tested do not appear to significantly influence (ceteris paribus) individual healthcare consumption. To begin with, no seasonal effect could be revealed, which is probably due to the fact that time elapsing between actual consumption and the receipt of the invoice is highly variable, which is likely to mask any seasonality effect. As to relative price of healthcare, it has not been found to significantly influence healthcare consumption. This is not entirely surprising considering that individuals only face part of the expenditure in most cases, because of their insurance coverage. Moreover, the time period covered by the study (2000–2005) might be too short to reveal any price effect, which should in fact be interpreted as a healthcare cost effect.

As to the model used, it appears that a Box-Cox transformation improves the fit compared with a log-transformation, while the correlation between selection and conditional consumption appears to be unnecessary, giving rise to a two-part model.

Finally, it may be remarked that total healthcare consumption is an aggregate of expenditures that are very different in nature. Further insight may be gained by estimating separate models for each type of healthcare, at least for drugs, physicians, and hospital care. This might raise new econometric issues, though. Indeed, the proportion of zero expenditures will be larger for those models, especially for hospital care. Moreover, consumptions for the different categories are correlated (except maybe for dental care), which should be addressed. Another development would be to introduce health insurance coverage into the model.

In fact, it will be interesting to investigate if the information provided by the SHIES makes it possible to tackle issues like moral hazard through the incentive effect of deductibles and co-payments. If so, this data source would constitute a very rich alternative to insurance data, which contain no individual characteristics other than age and gender, and are not representative of individuals living in Swiss private households.

References

Blough, David K., Carolyn W. Madden and Mark C. Hornbrook (1999), "Modeling Risk Using Generalized Linear Models", *Journal of Health Economics*, 18, pp. 153–171.

Box, George E. P. and David R. Cox (1964), "An Analysis of Transformations", *Journal of the Royal Statistical Society*, 26(2), pp. 211–252.

Chaze, Jean-Paul (2005), "Assessing Household Health Expenditure with Box-Cox Censoring Models", *Health Economics 2005*, 14, pp. 893–907.

Duan, Naihua, Willard G. Manning, Carl N. Morris and Joseph P. Newhouse (1984), "Choosing Between the Sample-Selection Model and the Multi-Part Model", *Journal of Business and Economic Statistics*, 2(3), pp. 283–289.

Gardiol, Lucien, Pierre-Yves Geoffard and Chantal Grandchamp (2005), "Separating Selection and Incentive Effects in Health Insurance", Working Paper n. 2005-38, Paris-Jourdan Science Economiques.

Gerfin, Michael and Martin Schellhorn (2006), "Nonparametric Bounds on the Effect of Deductibles in Health Care Insurance on Doctor Visits – Swiss Evidence", *Health Economics*, 15, pp. 1011–1020.

Holly, Alberto, Lucien Gardiol, Gianfranco Domenighetti and Brigitte Bisig (1998), "An Econometric Model of Health Care Utilization and Health Insurance in Switzerland", *European Economic Review*, 42, pp. 513–522.

Jones, Andrew M. (2000), "Health Econometrics", in: Cuyler A., Newhouse J. (Eds), *Handbook of Health Economics*, Amsterdam, vol. 1, pp. 265–344.

Jones, Andrew M. and Steven T. Yen (2000), "A Box-Cox Double-Hurdle Model", *Manchester School*, 68(2), pp. 203–221.

Lankford, R. Hamilton and James H. Wyckoff (1991), "Modeling Charitable Giving Using a Box-Cox Standard Tobit Model", *The Review of Economics and Statistics*, 73(3), pp. 460–470.

Lehmann, Hansjörg and Peter Zweifel (2004), "Innovation and Risk Selection in Deregulated Social Health Insurance", *Journal of Health Economics*, 23, pp. 997–1012.

Manning, Willard G., Anirban Basu and John Mullahy (2005), "Generalized Modeling Approaches to Risk Adjustment of Skewed Outcomes Data", *Journal of Health Economics*, 24, pp. 465–488.

Schellhorn, Martin (2001a), "The Effect of Variable Health Insurance Deductibles on the Demand for Physician Visits", *Health Economics*, 10, pp. 441–456.

Schellhorn, Martin (2001b), "A Comparison of Alternative Methods to Model Endogeneity in Count Models. An Application to the Demand for Health Care and Health Insurance Choice", Research Paper No. 40, SEDAP.

Schellhorn, Martin (2004), "Wählbare Selbstbehalte in der Krankenversicherung der Schweiz: Nachfragesteuerung oder Selektion?", Diskussionsschriften 04-02, Universität Bern.

Werblow, Andreas (2002), "Alles nur Selektion? Der Einfluss von Selbstbehalten in der Gesetzlichen Krankenversicherung", *Vierteljahrshefte zur Wirtschaftsforschung*, 71, pp. 427–436.

Yen, Steven T. (1993), "Working Wives and Food Away from Home: The Box-Cox Double-Hurdle Model", *American Journal of Agricultural Economics*, 75(4), pp. 884–895.

Yen, Steven T., Lynn E. Dellenbarger and Alvin R. Schupp (1995), "Determinants of Participation and Consumption: The Case of Crawfish in South Louisiana", *Journal of Agricultural and Applied Economics*, 27(1), pp. 253–262.

SUMMARY

The explanatory factors of individual healthcare consumption are studied by means of healthcare expenditures from the 2000–2005 Swiss Household Income and Expenditure Survey (SHIES). In order to tackle the issues of large number of null expenditures and skewed distribution of positive outcomes, the family of Box-Cox censoring models (Chaze, 2005) is applied. The results show that the use of SHIES data makes it possible to reveal many important factors of individual healthcare consumption, and that the role played by healthcare supply density variables is consistent with the theory of induced demand.