

## INCENTIVE EFFECTS IN THE DEMAND FOR HEALTH CARE: A BIVARIATE PANEL COUNT DATA ESTIMATION

REGINA T. RIPHAWN,<sup>a,\*</sup> ACHIM WAMBACH<sup>b,d</sup> AND ANDREAS MILLION<sup>a,c</sup>

<sup>a</sup> University of Munich, Germany

<sup>b</sup> CEPR, London, UK

<sup>c</sup> IZA, Bonn, Germany

<sup>d</sup> University of Erlangen-Nuremberg, Germany

### SUMMARY

This paper contributes in three dimensions to the literature on health care demand. First, it features the first application of a bivariate random effects estimator in a count data setting, to permit the efficient estimation of this type of model with panel data. Second, it provides an innovative test of adverse selection and confirms that high-risk individuals are more likely to acquire supplemental add-on insurance. Third, the estimations yield that in accordance with the theory of moral hazard, we observe a much lower frequency of doctor visits among the self-employed, and among mothers of small children. Copyright © 2002 John Wiley & Sons, Ltd.

### 1. INTRODUCTION

Since health care expenditures are substantial and rising in most industrialized countries it is important to understand their determinants.<sup>1</sup> Health care demand and its interaction with insurance choice are central factors in the development of health care expenditures. Since the relation between individuals and insurance providers is characterized by informational asymmetries, adverse selection and moral hazard problems are likely to affect the demand for services as well as the choice of insurance. This study contributes to the literature on health care demand in three dimensions. First, we introduce a new econometric technique to adequately model the processes generating demand for health care services. Second, we empirically evaluate adverse selection effects in the demand for health care, and third, we measure effects of moral hazard and discuss their policy relevance for the case of the German health care system.

The dependent variables of our analyses are integer valued count data. The application of count data estimation methods is well established in the literature on health care demand.<sup>2</sup> Pohlmeier and Ulrich (1995) as well as Cameron *et al.* (1988) use cross-sectional data with multiple outcomes per observation. A shortcoming of their econometric methods is the lack of controls for unobserved person specific heterogeneity. This type of heterogeneity is important in the analysis of health outcomes for two reasons. First, behavioural idiosyncracies such as the

---

\* Correspondence to: Regina T. Riphahn, WWZ-University of Basel, Petersgraben 51, CH-4003, Basel, Switzerland. E-mail: Regina.Riphahn@unibas.ch

<sup>1</sup> The share of German public and private health insurance expenditures in GDP increased from 4.2% in 1970, to 6.5% in 1990 and to 8.02% in 1997 (IW, various years). The share of health-related expenditures in the US budget increased from 11.3% to 19.2% between 1985 and 1995 (STBA, 1997).

<sup>2</sup> Cameron *et al.* (1988) apply a negative binomial (NEGBIN) model to their cross-section data, Geil *et al.* (1997) account for the panel nature of their data using a random effects NEGBIN estimator, and Pohlmeier and Ulrich (1995) estimate a NEGBIN hurdle model to reflect the two-part nature of the contact and intensity decision in their model of physician visits.

propensity to seek health care given a certain ailment might differ in unobservable but systematic ways between individuals. Second, individual genetic frailty and morbidity may generate significant yet unobservable influences on health care demand. An estimation approach which allows one to control for these unobservables improves the efficiency of the parameter estimates.

Geil *et al.* (1997) found that controlling for random effects in their one-equation count data model significantly improved the goodness of fit. We extend this approach to a bivariate count data model, which is new to this literature. In particular we specify a Poisson estimator with a lognormally distributed random effect for the bivariate count process. We jointly model individual demand for physician and for hospital visits based on data taken from the German Socioeconomic Panel (GSOEP, 1984–1995). While random effect procedures have been used in count data models before, this is the first application of a bivariate random effects panel estimator in a count data setting.

Our second contribution is to provide an empirical test for the existence of adverse selection effects in health care demand. Adverse selection describes a situation where the insured has better information about his risk type than the insurance, and he uses this information when buying insurance. As argued by Chiappori and Salani (2000), one of the fairly robust results in this context is that agents with higher risk probability choose contracts with higher coverage. The literature on health care demand has produced little evidence for the existence of adverse selection effects (for a survey see Marquis, 1992). Much of this literature is based on hypothetical experiments, where individuals are asked “what would you do if...”.<sup>3</sup> While the results indicate possible response patterns, they do not provide conclusive evidence on actual behaviours. After all, responses may differ between experimental decision frameworks and real-life health care decisions, a fact which is pointed out even by the researchers using experimental data (e.g. Marquis, 1992).<sup>4</sup> In contrast to the experiment-based literature our paper uses representative survey data on actual individual behaviours to investigate the relevance of adverse selection effects. Like Cameron *et al.* (1988) we test whether the choice of an insurance is correlated with and endogenous to subsequent health care demand.

In contrast to the adverse selection issue a number of studies have been devoted to the analysis of moral hazard, i.e. whether the demand for medical care of insured individuals exceeds that of identical uninsured persons. The notion of (*ex-post*) moral hazard describes the situation where the insured, due to not having to pay for medical expenses, consumes more than he would by paying the full price. One immediate consequence is that the demand for medical care of fully insured individuals exceeds that of identical individuals whose contracts specify larger copayments. Note that the costs of treatment to the patient may also be non-monetary. Using data from the RAND Health Insurance Experiment in the United States Manning *et al.* (1987) find substantial moral hazard effects. Geil *et al.* (1997) investigate the demand for hospital visits using German data and Cameron *et al.* (1988) study a variety of health care outcomes using Australian data. While the latter study confirms moral hazard effects in the utilization of some health care services, the former concludes that ‘insurance coverage does not play an important role for the hospitalization decision’.

We argue that the result of Geil *et al.* (1997) may be driven by their choice of a health care dimension which is particularly insensitive to economic incentives. Once one allows for a different patient willingness to pay for hospital versus doctor visits, theory suggests that the demand for

<sup>3</sup> For example, Marquis and Phelps (1987), van de Ven and van Vliet (1995); for an exception see Altman *et al.* (1998).

<sup>4</sup> For a discussion of the problems in establishing empirical evidence for adverse selection, see van de Ven and van Vliet (1995).

these services will respond differently to given incentives such as deductibles. This difference not only affects the appropriate formulation of an empirical model for health care demand, it is also crucial for the evaluation of incentive mechanisms set by public health policy.

Recent reforms in German public health insurance regulations increased co-payments on hospital expenditures without providing disincentives for doctor visits (Müller, 1995). Such a policy is a reasonable incentive mechanism only if the demand for hospital visits is more elastic than the demand for physician visits. Otherwise behavioural effects might be very small. This study compares the sensitivity of these dimensions of health care demand to economic incentives.

In the next section we briefly introduce institutional features of the German health care system. Section 3 discusses our hypotheses with respect to the incentive effects governing demand for hospital and doctor visits. Our data and the newly developed estimation method are described in Sections 4 and 5, before Section 6 of the paper reviews the results and Section 7 concludes.

## 2. MAIN INSTITUTIONAL FEATURES OF THE GERMAN HEALTH INSURANCE SYSTEM

The German health insurance system is characterized by almost complete coverage of the population. Health insurance is mandatory for all individuals with earnings below a cutoff value (DM 6450 gross monthly earnings in 2000, about \$3300). Exceptions are civil servants and the self-employed. These as well as persons with earnings above the cutoff can choose to remain uninsured, to voluntarily join one of the mandatory health insurances, a private insurance, or a combination of both. Currently about 90% of the population is insured in mandatory insurances and only about 0.1% of the population remains uninsured (STBA, 1997).

Typically individuals are insured in the health insurance chosen by their employer, with the retired and unemployed covered by the health insurance of their former employer. Family members of income earners are coinsured at no additional cost in the mandatory insurance scheme, while in private health insurances individual insurance has to be purchased for each family member.

Mandatory health insurances are mainly financed by payroll taxes that are split equally between the employer and the employee and are based on wage and salary incomes. The payroll tax rate differs between insurance funds, since each calculates the required rates dividing expected expenditures by the total tax base of its members. Prior to 1997 it was not possible to change health insurances. Thus health risks were pooled among the members of each insurance ('community rating'). While contribution rates differed somewhat across health insurances the benefits provided were generally very similar. In public insurances copayments and deductibles are limited to a few items such as glasses or prescription drugs. Copayments for hospital stays were introduced in 1989 and raised by 1995 from originally DM 10 to DM 12 per night for the first two weeks of in-patient care.

In contrast to public insurances private health insurances can condition their premia on individual health status, sex, and age. Many of the private insurances offer contracts which provide a rebate if no service is taken up during the calendar year which until recently was not possible for mandatory public insurances. In addition, most private insurers offer higher quality service than public insurers with advantages such as single or double rooms in hospitals, or treatment by the most experienced physicians. Thus it is not *a priori* clear whether a private insurance leads to more or less health care consumption.

Add-on insurance can be purchased in addition to pure public or private insurance. People who have public insurance might take up add-on private insurance which covers extra costs

like double rooms in hospitals or reimbursements for copayments on glasses.<sup>5</sup> In contrast to compulsory public insurance add-on insurance is optional and comes at a minimum of about DM 30 (about USD 17) per month, where premia vary substantially depending on the range of additional coverage.

### 3. DEMAND FOR HEALTH SERVICES AND HEALTH INSURANCE

The standard model of demand for health care is Grossman's (1972), where individuals invest in health capital and demand health services similar to how they invest in human capital. However, in contrast to human capital, health capital is in addition valued as a consumption good, as sick days enter the utility negatively. There is a stock of health capital, which depreciates over time:  $H_{i+1} - H_i = I_i - \delta_i H_i$ .  $\delta_i$  is the rate of depreciation during period  $i$ .  $I_i$  is the investment in health capital which itself depends on medical care, time invested in health, and the stock of human capital. A shortcoming of this model is that the risk aspect of the demand for health services is not included, so that the insurance decision cannot be modelled explicitly.

This aspect is taken up by Cameron *et al.* (1988) which we take as the basis for our analysis. In their model the expected utility of an individual for a given insurance policy can be written as:

$$E[U] = \int_s U(C_0, C_1, H(e, s)|A, B) d\pi(s|A) + w \quad (1)$$

where  $s$  denotes the state of the world,  $C_i$  is consumption in period  $i$ ,  $H$  denotes health,  $e$  is a  $k$ -dimensional vector of health services,  $A$  are consumer attributes,  $B$  are attributes of the insurance policy, and  $w$  is an error term.

The individual chooses insurance in a dynamic programming approach, where first expected utility is optimized under each insurance option with respect to  $C_0$ ,  $C_1$  and  $e$ , and subject to wealth constraints. Then the actual insurance policy is determined based on a comparison of the conditional expected utility for each insurance option. It is assumed that the choice of insurance determines the price that consumers pay for health services.

For the empirical model the utility function is specified assuming that

$$U(\cdot) = C_0 C_1^{\alpha+1} H(e|s, A, B)^{\sigma+1}$$

with  $H = \prod_k e_k^{\alpha_k(s, A, B)}$  which can be interpreted as a health production function.  $\sigma$  represents the constant coefficient of relative risk aversion.

Solving the optimization problem for a given choice of insurance leads to an equation for the demand of health services which depends on the state of the world and the insurance choice.

<sup>5</sup> Our survey of insurance firms yielded that beyond the typical indicators such as age and sex, some firms used information on the type of employment to calculate the premium. However, by law the insured are required to indicate whether they have any further relevant information concerning their health status, even if the insurer does not ask explicitly for this information. We have no information as to how far this law is applied. Thus we are unable to distinguish adverse selection in our context from the notion of 'essential adverse selection' proposed by van de Ven and van Vliet (1995). 'Essential adverse selection' refers to the residual adverse selection, after insurers have conditioned their contracts on all available information.

To obtain a tractable form for the estimation equation further assumptions are needed regarding the distribution of illness states. Cameron *et al.* (1988) obtain that unconditional of insurance choice the demand for the  $k$ th medical service can be written as:

$$E[e_k(s)] = \exp \left( Z' \beta_k + \sum_{j=1}^J \eta_{jk} D_j + \varepsilon_k \right) \quad (2)$$

where  $Z$  is a vector of covariates and  $D_j$  are dummy variables for the  $j$ th insurance form.

There are two implications of this model which are the main focus of the following analysis. First, from the derived demand equation one obtains that if the price of the medical service is lower, more of it will be demanded. Thus one would expect that  $\eta_{jk}$  is larger for those policies  $j$  which are more generous. This is the well-known moral hazard effect of health insurance. Second, the model by Cameron *et al.* points towards the endogeneity of the insurance choice. Consumers will calculate their expected utility under the different policies and choose that insurer which provides the highest utility. Although Cameron *et al.* do not state this explicitly, from other models (e.g. Rothschild and Stiglitz, 1976) it is well known that high-risk types are expected to buy more comprehensive coverage.

Both of these issues are investigated in turn. We proceed as Cameron *et al.* by first estimating the unconditional demand equations to see to what extent the choice of insurance determines the demand for health care. In a second step we analyse whether insurance choice can be considered as endogenous.

#### 4. DESCRIPTION OF THE DATA

The data are taken from the first twelve annual waves (1984 through 1995) of the German Socioeconomic Panel (GSOEP) which surveys a representative sample of East and West German households (Wagner *et al.*, 1993). The data provide detailed information on the utilization of health care facilities, characteristics of current employment, and the insurance schemes under which individuals are covered. Following Geil *et al.* (1997) we restrict our sample to individuals aged 25 through 65,<sup>6</sup> from the West German subsample, and of German nationality. After dropping observations with missing values on key variables the sample contained 3691 male and 3689 female individuals which make up a sample of 14,243 male and 13,794 female person-year observations.<sup>7</sup>

Our dependent variables are the number of visits to a doctor within the last quarter prior to the survey, and the number of inpatient hospital visits (with at least one night spent in the hospital) within a given calendar year. The information on hospital visits for a year  $t$  is gathered only in the survey of year  $t + 1$ . This information is matched here with the number of physician visits observed for year  $t$ . Since the relevant questions were not asked in the surveys of 1990 and 1993 we can use matched dependent variable pairs for the years 1984–1988, 1991, and 1994.<sup>8</sup> Both dependent variables are integer valued. Descriptive statistics are presented in Table I by gender. Hospital visits are highly skewed at zero, and the average number of physician visits amounts

<sup>6</sup> By this age restriction we exclude retirees and individuals who are still covered by their parents' health insurance.

<sup>7</sup> Observations of women who gave birth within the period of six months before and after a given calendar year were omitted from the sample for the relevant years.

<sup>8</sup> Due to the omitted questions we do not observe hospital visits for 1989 and 1992, and we lack information on physician visits for 1990 and 1993.

Table I. Dependent variables—descriptive statistics

Value	(Share of total observations, %)			
	Hospital visits		Doctor visits	
	Male	Female	Male	Female
0	92.2	90.2	44.0	29.5
1	6.2	7.9	13.8	13.2
2	1.1	1.3	11.6	13.4
3	0.2	0.3	8.5	11.5
4–9	0.1	0.3	15.4	21.8
10	0.2	0.1	6.7	10.6
Mean	0.128	0.150	2.63	3.79
Std dev.	0.930	0.831	5.21	6.11
Median	0	0	1	2
<i>N</i>	14,243	13,083	14,243	13,083

Source: German Socioeconomic Panel (1984–1995).

to 2.6 and 3.8 for men and women, respectively. The statistically significant positive correlation between the two health care utilization measures takes on values of 0.15 for the male and 0.13 for the female sample.

To evaluate differences between insurance types we generated two indicator variables. The first describes whether an individual is covered by the public sector of the German health insurance system, versus being privately insured. This indicator groups those who are obliged to take up mandatory health insurance, with those voluntarily in the public health insurance. The latter group consists mostly of employees with incomes above the earnings cutoff value, and makes up about 15% of those under public insurance coverage. The group of individuals who are privately insured combines civil servants, and those who indicated a private insurance. They make up about 14% in the male and 9% in the female subsample. The second insurance indicator describes whether an individual who is covered by the public insurance system purchased add-on insurance, as described in Section 2 above. This applies for only about 2% of the publicly insured subsamples.

For descriptive statistics on these and other explanatory variables see Table II. Clear differences between the male and female subsamples are found with respect to health indicators, where men are more frequently handicapped but have, on average, higher health satisfaction than women. On average, men have slightly more years of education. Also, 85% of the male sample is employed compared to 49% of females, and among employed men the fractions who are self-employed, civil servants, or blue collar workers are higher.

Table III presents mean outcomes of the independent variables for different subsamples. It is apparent that health care utilization is higher in the group of publicly health insured individuals than among those who purchased private insurance. We do not detect higher utilization of physicians among those who purchase add-on insurance. That group, however, has an above average frequency of hospital visits. The difference appears to be stronger for the male than the female subsample. Table III further indicates higher health care utilization for the aged and handicapped. Self-employed individuals seem to visit the doctor by about one third less for both subsamples while hospitalization rates do not differ as much. Being married is correlated with a higher number of physician visits for women and a smaller number for men while both sexes visit the hospital less

Table II. List of variables

Variable	Description	Males <sup>a</sup>		Females <sup>a</sup>	
<i>doc</i>	Number of doctor visits in last 3 months	2.626	(5.21)	3.791	(6.11)
<i>hosp</i>	Number of hospital visits last year	0.128	(0.93)	0.150	(0.83)
<i>age</i>	Age	42.653	(11.27)	44.476	(11.32)
<i>health</i>	Health satisfaction coded 0 (low)–10 (high)	6.924	(2.25)	6.634	(2.33)
<i>handicap</i>	Person is handicapped (0/1)	0.227	(0.42)	0.200	(0.40)
<i>h_degree</i>	Percentage degree of handicap	8.134	(20.33)	5.791	(17.96)
<i>married</i>	Person is married (0/1)	0.765	(0.42)	0.752	(0.43)
<i>schooling</i>	Years of schooling	11.729	(2.44)	10.876	(2.11)
<i>hhincome</i>	Monthly household net income ( $\times 10^{-3}$ )	3.591	(1.74)	3.445	(1.80)
<i>child</i>	Child(ren) below age 16 in household (0/1)	0.413	(0.49)	0.392	(0.49)
<i>self-employed</i>	Person is self-employed (0/1)	0.086	(0.28)	0.037	(0.19)
<i>civil servant</i>	Person is civil servant (0/1)	0.118	(0.32)	0.028	(0.16)
<i>blue collar</i>	Person is blue collar worker (0/1)	0.340	(0.47)	0.139	(0.35)
<i>employed</i>	Person is employed (0/1)	0.850	(0.36)	0.488	(0.50)
<i>public ins</i>	Person is insured in public health insurance (0/1)	0.861	(0.35)	0.913	(0.28)
<i>add on ins</i>	Person is insured in add-on insurance (0/1)	0.018	(0.13)	0.020	(0.14)
<i>1985</i>	Year = 1985(0/1)	0.139	(0.35)	0.139	(0.35)
<i>1986</i>	Year = 1986(0/1)	0.138	(0.34)	0.139	(0.35)
<i>1987</i>	Year = 1987(0/1)	0.134	(0.34)	0.134	(0.34)
<i>1988</i>	Year = 1988(0/1)	0.162	(0.37)	0.166	(0.37)
<i>1991</i>	Year = 1991(0/1)	0.158	(0.36)	0.160	(0.37)
<i>1994</i>	Year = 1994(0/1)	0.127	(0.33)	0.120	(0.32)
<i>N</i>	number of observations	14,243		13,083	

<sup>a</sup> means, standard deviations in parentheses.

Source: German Socioeconomic Panel (1984–1995).

often when married. With children in the household health care utilization generally increases. Household net income appears to be negatively correlated with health care utilization, except for women's hospital visits.

## 5. THE BIVARIATE PANEL COUNT DATA MODEL

To study the determinants of health care utilization we use a variant of a new class of count data models called LogNormalP. This class modifies the Poisson model and was introduced by Million (1998). Consider the basic framework first.

In a Poisson model it is assumed that the dependent variable,  $y_i$ , is Poisson distributed. This distribution has only one parameter  $\lambda_i$ , the so-called intensity. Covariates are introduced using an exponential parameterization of  $\lambda_i$  such that the expected value of  $y_i$  is a function of the covariates

Table III. Mean health care utilization by selected characteristics

	Doctor visits		Hospital visits	
	Male	Female	Male	Female
Total	2.63	3.82	0.128	0.174
Public insurance				
Yes	2.74	3.92	0.133	0.176
No	1.90	2.77	0.099	0.155
Add-on insurance				
Yes	2.50	3.63	0.176	0.187
No	2.63	3.82	0.127	0.174
Age				
25–35	1.95	3.21	0.107	0.235
35–45	1.92	3.30	0.111	0.144
45–55	2.90	3.95	0.141	0.136
55–65	4.28	4.98	0.166	0.175
Handicap				
Yes	4.01	5.31	0.178	0.244
No	2.22	3.45	0.113	0.157
Self-employed				
Yes	1.88	2.72	0.118	0.106
No	2.70	3.86	0.129	0.176
Married				
Yes	2.34	4.04	0.118	0.164
No	2.72	3.74	0.131	0.177
Children in household				
Yes	2.98	4.27	0.132	0.168
No	2.13	3.16	0.122	0.182
Household net income				
<2400	3.20	4.22	0.142	0.165
2400–3200	2.78	3.69	0.141	0.167
3200–4300	2.43	3.63	0.129	0.185
>4300	2.20	3.43	0.102	0.178

Source: German Socioeconomic Panel (1984–1995).

$x_i$  and coefficients  $\beta$ :

$$y_i \sim Po(\lambda_i) : p(y_i = a) = \frac{e^{-\lambda_i} \lambda_i^a}{a!} \quad (3)$$

$$E(y_i) = \text{Var}(y_i) = \lambda_i = e^{\beta' x_i} \quad (4)$$

The Poisson model is restrictive, first, because it demands equality of mean and variance which in many applications is inappropriate, and second, because it does not permit generalizations to situations with correlated outcomes. These problems have led to a large number of new models (see Cameron and Trivedi, 1998). Most popular is the NEGBIN model with its extensions for panel data. It was introduced to econometrics by Hausman *et al.* (1984) and Cameron and Trivedi (1986). In the NEGBIN model the deterministic nature of the link between intensity  $\lambda_i$  and covariates  $x_i$  is relaxed by consideration of a gamma distributed random error. Conditional on it the count  $y_i$  is again Poisson distributed. Since the error is not observable, one has to integrate it out in order



to obtain unconditional expressions for maximum likelihood estimation. The main characteristics of the NEGBIN model are overdispersion, i.e. conditional on the covariates the variance exceeds the mean, and a likelihood function which is easy to implement. However, the assumption of a gamma distributed error has no statistical justification and is motivated purely by mathematical convenience.

The restrictions of the basic Poisson model may more appropriately be overcome by an alternative set of distributional assumptions. Preston (1948) proposed a lognormal error distribution which leads to a Poisson lognormal distribution. This model has the following form:

$$p(y_i = a | \lambda_i) = \frac{e^{-\lambda_i} \lambda_i^a}{a!} \quad (5)$$

$$\ln(\lambda_i) \sim N(\beta'x_i, \sigma) \quad (6)$$

The likelihood function reveals immediately why this model has not been as popular as NEGBIN in the past:

$$L(\beta, \sigma) = \prod_{i=1}^n \int_0^\infty \frac{e^{-\lambda_i} \lambda_i^{y_i}}{y_i!} \frac{1}{\sqrt{2\pi\sigma\lambda_i}} e^{-\frac{(\ln \lambda_i - \beta'x_i)^2}{2\sigma^2}} d\lambda_i \quad (7)$$

As in the NEGBIN case the error has to be integrated out. But here the resulting integrals cannot be solved analytically and thus the likelihood function contains integrals, which can only be solved using numerical integration procedures. Such procedures have been applied in econometrics before and can now be utilized at drastically reduced evaluation times due to the progress in computer processing power. Therefore assuming lognormally distributed errors is now a feasible alternative to the NEGBIN model. The advantage of using the assumption of lognormality, which we label the LogNormalP model is that it provides a flexible framework for fully parametric count data models. To see this we reformulate statement (6) such that:

$$\ln(\lambda_i) = \beta'x_i + \varepsilon_i \quad (8)$$

where  $\varepsilon_i$  is a normally distributed error term. The right-hand side of equation (8) is familiar from linear models or from non-linear models with a latent variable. This similarity is the key to building complex count data models. Popular concepts such as the random effects panel model can now be adopted by adjusting the correlation structure of  $\varepsilon_i$ . Also, new models can easily be built by customizing the correlation between different errors  $\varepsilon_i$  and  $\varepsilon_j$ . This is the main advantage of a lognormal over a gamma distribution which provides no options for direct generalizations to multivariate counterparts.

A statistical reason for the normality assumption of  $\varepsilon_i$  is pointed out by Winkelmann (2000): If an error captures the effect of omitted regressors we can establish normality by the central limit theorem. For further comparisons of the gamma and normal distribution assumptions see Million (1998).

Given the observed correlation between the two dimensions of health care demand, i.e. doctor and hospital visits, it appears plausible to formulate an estimation model which accounts for possible correlation in these two outcomes' unobservable determinants. In our bivariate panel model we assume that every person  $i$  has  $g$  different outcome observations  $y$  (i.e. doctor and hospital visits) in any period  $t$ .

$$y_{itg} \sim Po(\lambda_{itg}), \quad g = 1, 2 \quad (9)$$

$$\ln(\lambda_{it1}) = \beta'_1 x_{it1} + u_{i1} + \varepsilon_{it1} \quad (10)$$

$$\ln(\lambda_{it2}) = \beta'_2 x_{it2} + u_{i2} + \varepsilon_{it2} \quad (11)$$

$$(\varepsilon_{it1}, \varepsilon_{it2}) \sim N_2(0, 0, \sigma_{\varepsilon 1}^2, \sigma_{\varepsilon 2}^2, \rho) \quad (12)$$

$$u_{i1} \sim N(0, \sigma_{u1}^2) \quad (13)$$

$$u_{i2} \sim N(0, \sigma_{u2}^2) \quad (14)$$

$$E[\varepsilon_{itg} u_{jgh}] = 0 \quad \forall i, t, g, j, h \quad (15)$$

$$E[\varepsilon_{itg} \varepsilon_{jsh}] = 0 \quad \text{if } t \neq s \vee i \neq j \vee g \neq h \quad (16)$$

$$E[u_{ig} u_{jgh}] = 0 \quad \text{if } i \neq j \vee g \neq h \quad (17)$$

Then equation (9) describes the distributional assumption for  $y$ , equations (10) and (11) represent the estimation models which are characterized by a dual error structure. The random errors  $\varepsilon$  follow a bivariate normal distribution while the outcome specific error terms  $u$  represent individual specific unobserved heterogeneity which is constant over time. Statements (12) to (17) describe the distributional assumptions.

Thus, in this model there are two possible sources of correlation between the unobservable determinants of the two dimensions of demand for health care. First, correlation across time between observations of any given person is captured by the outcome specific error component  $u$ . This error component captures an individual's unobserved latent propensity to demand a certain type of health care. Examples are the repeated need for dental treatments for individuals with bad teeth, or dialysis treatment in a hospital for those with kidney problems. In principle such diagnoses are observable, but our data does not provide the necessary information. Second, the error terms  $\varepsilon_{itg}$  control for shocks which in any period may affect both outcomes. Such correlated period shocks can result from flu epidemics which affect all individuals jointly and cause a correlation in  $t$  across all  $i$ . Alternatively a given individual ailment may cause a person to first consult a physician before going to the hospital. In either case  $\varepsilon_{it1}$  and  $\varepsilon_{it2}$  may be correlated. In contrast to the panel correlations captured in the distributions of  $u_{i1}$  and  $u_{i2}$  their correlation can also be negative, for example when outpatient treatment temporarily substitutes inpatient treatment. Besides these possible correlations one might in principle control for general period shocks across individuals, or for general individual factors across equations. However, since the identification of such complex error term structures becomes increasingly difficult, we focus on the sufficiently flexible approach presented above.<sup>9</sup>

The likelihood function for this bivariate panel estimator is complicated, because it involves analytically intractable integrals. The joint errors for given individuals  $u$  and across equations  $\varepsilon$  have to be integrated out numerically. We solved the problem by Gaussian quadrature procedures, where the distribution of the individual-specific unobservables was integrated out in an outer integral using a Gauss–Hermite approximation, and the distribution of the cross-equation errors was integrated in an inner integral following a modified Gauss–Legendre approach. This modified Legendre procedure dominated the available alternatives (Hermite and Laguerre integration) in

<sup>9</sup> See Winkelmann (2000) for a survey of recent developments in the literature applying time series estimation approaches to count data models. Due to the interrupted nature and infrequency of our annual observations, these methods are hardly applicable here.

terms of computing time and bias. The likelihood function was optimized using a quasi-Newton procedure (for further details on the estimation procedure see Million, 1998).

Before turning next to a discussion of estimation results it may be helpful to point out how the coefficients of count data models can be interpreted. In count data models the magnitudes of different factors' influences are easily calculated. Two standard cases have to be discussed. If  $x_c$  is a continuous variable then  $\beta_c$  is a semi-elasticity. An increase of  $dx_c = 0.01$  changes the expectation of  $y$  by  $\beta_c$  %:

$$\frac{\partial E(y|x)}{\partial x_c} = \beta_c E(y|x) \Rightarrow \beta_c = \frac{\partial E(y|x)}{\partial x_c} \frac{1}{E(y|x)} \quad (18)$$

This marginal view is inappropriate for binary explanatory variables,  $x_b$ . In this case a direct comparison of both values is more informative. For small values,  $\beta_b$  is a measure for the proportional change of the expectation shifting  $x_b$  from 0 to 1.

$$\frac{E(y|x_b = 1, x)}{E(y|x_b = 0, x)} = \frac{e^{\beta_b + \sum_{i \neq b} \beta_i x_i}}{e^{\sum_{i \neq b} \beta_i x_i}} = e^{\beta_b} \quad (19)$$

$$\text{For } |\beta_b| \ll 1 : e^{\beta_b} \approx 1 + \beta_b \quad (20)$$

## 6. RESULTS

### 6.1. Model Performance

To judge the merit of the bivariate panel approach against simpler estimation procedures we estimated both models. Table IV lists the results for the univariately estimated models without random effects. Comparing these results with those of the bivariate panel estimation in Table V yields various differences.<sup>10</sup> For several variables the statistical significance changes dramatically between the estimations, e.g. *age* and *age squared* are individually highly significant for hospital trips of men in Table V and insignificant in Table IV. The opposite holds for other variables such as marital status or children in the equation for males' doctor visits and among others for marital and handicap status or blue collar employment for females' doctor visits. It is therefore crucial to use the better suited estimation approach.

In order to test for the appropriate model one can use a standard likelihood ratio test, because the bivariate panel model nests the univariate cross-sectional model:

$$LRT_{\text{men}} = 2 \times (-31,227.1 + 27,411.4 + 4,615.1) = 1,598.8 \quad (21)$$

$$LRT_{\text{women}} = 2 \times (-34,391.8 + 30,213.4 + 4,998.1) = 1,639.4 \quad (22)$$

These test statistics lead us to reject the simple model, which imposes zero correlation in individual observations over time and across contemporary outcomes. The statistics confirm the high statistical significance of the structural parameters  $\sigma_u$  and  $\rho$  which is found for the individual parameters already in Table V.<sup>11</sup> Comparing estimation results for  $\sigma$  (Table IV) and  $\sigma_e$  (Table V) shows

<sup>10</sup> The specifications in Tables IV and V were estimated controlling for a set of year fixed effects. These coefficients are omitted in the tabulations to save space.

<sup>11</sup> The critical value for the likelihood ratio test at the 1% significance level with two degrees of freedom is 10.6.

Table IV. LogNormalP

	Males		Females	
	Doctor	Hospital	Doctor	Hospital
Constant	2.775** (10.69)	-1.047* (-1.70)	2.752** (12.40)	-1.368** (-2.39)
<i>age</i>	-0.057** (-5.18)	-0.017 (-0.68)	-0.044** (-4.83)	-0.033 (-1.33)
<i>age</i> <sup>2</sup> · 10 <sup>-3</sup>	0.773** (6.04)	0.221 (0.77)	0.531** (5.17)	0.218 (0.78)
<i>health</i>	-0.265** (-43.37)	-0.233** (-18.33)	-0.223** (-42.65)	-0.207** (-16.06)
<i>handicap</i>	-0.000 (-0.01)	-0.156 (-1.36)	0.122** (2.50)	0.077 (0.64)
<i>h_degree</i>	0.007** (7.59)	0.009** (4.71)	0.005** (5.71)	0.011** (5.77)
<i>married</i>	0.146** (4.13)	-0.078 (-0.86)	0.051* (1.77)	-0.060 (-0.78)
<i>schooling</i>	-0.023** (-3.26)	-0.041** (-2.41)	0.013* (1.89)	-0.021 (-1.29)
<i>hhincome</i>	-0.135 (-1.51)	0.281 (1.33)	-0.143* (-1.86)	0.491** (2.60)
<i>child</i>	-0.107** (-3.37)	0.070 (0.89)	-0.140** (-4.83)	0.015 (0.19)
<i>self-employed</i>	-0.268** (-4.70)	-0.032 (-0.26)	-0.301** (-4.19)	-0.091 (-0.48)
<i>civil servant</i>	0.045 (0.77)	-0.103 (-0.72)	-0.037 (-0.45)	0.302 (1.31)
<i>blue collar</i>	-0.013 (-0.35)	0.165** (1.97)	-0.095** (-2.67)	-0.342** (-3.11)
<i>employed</i>	0.005 (0.11)	-0.186* (-1.78)	0.044 (1.56)	0.009 (0.11)
<i>public ins</i>	0.073 (1.37)	-0.198 (-1.50)	0.074 (1.51)	0.305** (2.50)
<i>add on ins</i>	0.166 (1.64)	0.566** (2.81)	0.132 (1.59)	0.263 (1.26)
$\sigma$	1.267** (100.08)	1.529** (65.33)	1.062** (103.53)	1.518** (54.39)
<i>lnL</i>	-27,411.4	-4,615.1	-30,213.4	-4,998.1
<i>n</i>	14,243	14,243	13,083	13,083

*t*-values are given in parentheses.

Two-sided test: \*10% significance, \*\*5% significance.

Table V. Bivariate PanelLogNormalP

	Males		Females	
	Doctor	Hospital	Doctor	Hospital
Constant	2.563** (8.52)	−0.206 (−0.27)	2.423** (9.40)	−1.567** (−2.29)
<i>age</i>	−0.060** (−4.24)	−0.077** (−2.35)	−0.040** (−3.73)	−0.032 (−1.11)
<i>age</i> <sup>2</sup> · 10 <sup>−3</sup>	0.823** (5.10)	0.942** (2.49)	0.499** (4.15)	0.234 (0.73)
<i>health</i>	−0.237** (−33.07)	−0.243** (−16.03)	−0.191** (−35.81)	−0.196** (−13.41)
<i>handicap</i>	−0.029 (−0.63)	−0.086 (−0.73)	0.063 (1.44)	0.039 (0.34)
<i>h_degree</i>	0.007** (6.90)	0.008** (3.27)	0.004** (4.75)	0.010** (4.67)
<i>married</i>	0.085 (1.64)	−0.054 (−0.49)	0.009 (0.26)	−0.044 (−0.48)
<i>schooling</i>	−0.022** (−2.39)	−0.051** (−2.35)	0.014 (1.62)	−0.015 (−0.71)
<i>hhincome</i>	−0.090 (−0.91)	0.375 (1.55)	−0.107 (−1.31)	0.407* (1.92)
<i>child</i>	−0.059 (−1.44)	0.103 (1.08)	−0.117** (−3.58)	0.073 (0.82)
<i>self-employed</i>	−0.356** (−3.84)	−0.196 (−1.17)	−0.256** (−3.54)	−0.117 (−0.57)
<i>civil servant</i>	−0.011 (−0.15)	−0.086 (−0.46)	−0.069 (−0.64)	0.281 (1.06)
<i>blue collar</i>	−0.029 (−0.65)	0.173 (1.56)	−0.034 (−0.82)	−0.320** (−2.68)
<i>employed</i>	0.041 (0.79)	−0.026 (−0.19)	0.002 (0.07)	−0.014 (−0.16)
<i>public ins</i>	0.075 (1.19)	−0.136 (−0.85)	0.058 (1.05)	0.246 (1.63)
<i>add on ins</i>	0.090 (0.87)	0.549** (2.08)	0.096 (1.17)	0.219 (1.02)
$\sigma_\varepsilon$	0.996** (64.85)	1.244** (25.18)	0.822** (80.89)	1.053** (22.82)
$\sigma_u$	0.795** (34.75)	1.195** (19.90)	0.701** (42.23)	1.123** (21.88)
$\rho$	0.490** (14.27)		0.386** (9.94)	
<i>lnL</i>	−31,227.1		−34,391.8	
<i>n</i>	14,243		13,083	
time in h	17		14	

*t*-values are given in parentheses.

Two-sided test: \*10% significance, \*\*5% significance.

the reason for this predominance: Decomposing the equation error of the simple model reduces variation of the observation specific error term. Therefore it is highly recommendable to use the model for correlated counts in appropriate data situations.<sup>12</sup>

## 6.2. Adverse Selection

The above discussion of adverse selection in the German health insurance system yielded that those individuals who foresee higher health expenditures can be expected to purchase add-on insurance. For the empirical model this suggests two predictions. First, we should observe a positive correlation between holding add-on insurance and realized health care demand. Second, holding add-on insurance should be endogenous to subsequent health care demand.

An inspection of Table V yields that indeed all coefficients of the add-on insurance indicator in the health care demand equations for men and women are positive. The coefficients in the doctor visit equations are of small magnitudes and not statistically significant. Given the benefit packages of German add-on insurances which typically finance higher quality hospital care or cover co-payments for medications and medical aids, the imprecise effect on doctor visits is not surprising. The magnitudes of the coefficients in the hospital visit equations are much larger. Based on the approximation presented in equation (20), holding add-on insurance increases the expected number of hospital visits for men by 55 and for women by 22% which are large effects. Interestingly, only the coefficient for the male sample is statistically significant confirming the differential univariate effects already observed in Table III. Therefore this first evidence appears to corroborate the expected adverse selection effects.

To test our hypothesis regarding the endogeneity of add-on insurance to subsequent health care demand we apply the procedure used by Cameron *et al.* (1988). They predicted dichotomous insurance indicators using an instrumental variables procedure. Similarly, we apply a binary logit estimator to predict whether an individual purchases add-on insurance. The predicted values are then considered in a least squares regression of health care demand based on which a Hausman test can be applied. If the coefficient of the predicted value of add-on insurance is statistically significant in the models for health care demand, the hypothesis of exogeneity is rejected (see Hausman, 1978, p. 1260).

The prediction equations for add-on insurance were estimated separately for men and women using a logit estimator. They control for detailed indicators of individual age, human capital, and as the key instrument for the number of private health insurance firms in an individual's state of residence. The validity of the instrument was subjected to an overidentification test which yielded that the health insurance indicator as well as the categorical age variables were statistically significant in the model of health insurance, but insignificant when added to the demand for care equations. The results are presented in the Appendix.<sup>13</sup> Least squares regressions of doctor and

<sup>12</sup> In separate estimations we tested the effects of controlling for the panel nature of the data and the correlation of the outcome variables, separately. The results suggest that each of the two corrections yielded an improvement which is statistically significant at the 1% level. However, the likelihood effect of controlling for the panel nature of the data is much larger than that of controlling for the correlation between the dependent variables.

<sup>13</sup> In our sample of publicly and privately insured individuals only the former can choose to purchase add-on insurance. Therefore the endogeneity test applies only to this group. In addition, the sample of publicly insured consists of those individuals who are mandatorily insured, and those who can choose public over private insurance. The decision to purchase add-on insurance differs between these two groups; while the former take a dichotomous choice, the latter optimize over a bundle of options including private insurances, public insurance with and public insurance without add-on insurance.

Table VI. Ordinary least squares

	Males		Females	
	Doctor	Hospital	Doctor	Hospital
Predicted add-on insurance	5.715 (1.172)	3.809* (1.786)	-0.286 (0.119)	0.812* (1.728)
$R^2$	15.11	1.12	15.21	2.15
$N$	9,274	9,274	11,669	11,669

In parentheses  $t$ -values based on White-corrected standard errors.

Two-sided test: \*10% significance, \*\*5% significance.

Other covariates as in Table V, only public ins is excluded.

hospital visits then controlled for the variables as in Table V plus the predicted values of add-on insurance, omitting only the indicator variable for public insurance. Table VI presents the coefficients of these predicted values in the four least squares regressions. Exogeneity is rejected if the presented coefficient is statistically significantly different from zero. This is the case at the 10% significance level for the male and female hospital visits. These findings are robust to changes in the specification of the add-on insurance prediction equation. While the results cannot be interpreted as unambiguous proof of adverse selection into add-on insurance we consider them as suggestive evidence in that direction.<sup>14</sup>

This evidence compares well with other results. In the study of a Swiss firm which implemented a new health insurance scheme with distinct health insurance options, Ellis (1989) provides evidence of health expenditure differences greater than 500% between the group with the lowest deductible and that with the largest deductible. Holly *et al.* (1998) strongly reject the exogeneity of the insurance purchase to subsequent health care utilization and show that holding supplemental insurance significantly increases the probability that a person has at least one inpatient stay, conditional on having used some medical treatment. Altman *et al.* (1998) report that people switching plans (from HMO to indemnity) had expenditures that were 47% higher than those remaining in the HMO. The overall effect of adverse selection, i.e. the effect on the premia for the two plans is only about 1%.

### 6.3. Moral Hazard

Turning now to the incentives affecting demand for health care, moral hazard driven behaviour would imply that individuals with higher insurance coverage, and therefore lower opportunity cost, demand more health care. This has been tested numerous times in the literature.<sup>15</sup> However, the hypothesis could not be confirmed in panel data models for Germany so far.<sup>16</sup> Therefore we

Unfortunately we do not have enough observations in this latter group to analyse the determinants of the set of different outcomes. Therefore we focus on the former group of individuals who are mandatorily in the public insurance, which provides a homogenous sample of individuals with a comparable decision framework.

<sup>14</sup> In order to determine whether considering the potentially endogenous add-on variable in the model affects the coefficient estimates on other covariates we reran the model omitting the add-on indicator. Since the estimation results on the other coefficients are robust to this modification in the specification we are confident that the endogeneity of the insurance indicator does not affect our conclusions.

<sup>15</sup> See e.g. Long *et al.* (1998), Coulson *et al.* (1995), or for a classic Manning *et al.* (1987).

<sup>16</sup> In their cross-sectional analysis Pohlmeier and Ulrich (1995) find a significantly lower propensity to contact a general practitioner among those in private insurance, but a higher propensity to contact a specialized physician.

generalize the test for moral hazard behaviour applied by Geil *et al.* (1997) by considering different outcomes of health care demand. Our estimation results (see Table V) suggest that the insurance type, i.e. whether individuals are privately insured or covered by public health insurance, does not yield a statistically significant influence on the demand for health care. Neither the number of hospital nor of doctor visits are significantly affected by the health insurance indicator.<sup>17</sup> This confirms the results by Geil *et al.* (1997) and by Chiappori *et al.* (1998) who find that a change in relative prices has no influence on the demand for general practitioner or specialists office visits in France.

There are three estimation results which raise doubts about a conclusion of no moral hazard effects. First, we found that the presence of add-on insurance is correlated with higher demand for hospital visits. While the tests for endogeneity of this insurance choice suggest that the positive effect on actual demand is at least in part due to *ex ante* adverse selection, the possibility of *ex-post* moral hazard cannot be excluded. Second, and more interesting, is the negative effect of self-employment status on the demand for doctor visits. The coefficients of this variable are large and suggest that self-employed males (females) visit doctors about 35 (25)% less often than the not self-employed. The coefficients are statistically highly significant for both samples. It is important to know that at the time of our data, employees were fully insured for loss of wage incomes when ill. In contrast the self-employed visit the doctor without compensation, and therefore at large financial disincentives. The finding of significantly fewer doctor visits for this group therefore confirms incentive effects in the demand for health care. The third finding indicative of the relevance of incentive effects is the role children in the household play for female health care demand. Again, doctor and hospital visits are differently affected. And again demand for doctor visits responds significantly negatively to the increased opportunity cost of health care demand brought about by the presence of children.<sup>18</sup>

#### 6.4. Other Findings

Turning now to the interpretation of the results in Table V, we find a number of statistically significant determinants of health care demand. As expected, the indicator for high health satisfaction is negatively and the percentage degree of a handicap is significantly positively correlated with the demand for both doctor and hospital visits. Also, the quadratic age effect is statistically significant in all models but female hospital visits, confirming the descriptive statistics presented in Table III. Interestingly, male health care demand decreases with better education, confirming the results of other studies such as e.g. Geil *et al.* (1997) or Pohlmeier and Ulrich (1995). In contrast, female demand for health care does not appear to respond to schooling which may be due to the larger out of the labour force sample in this group for whom schooling differences may matter less.

### 7. CONCLUSIONS

Following up on recent reforms of the public insurance system which increased co-payments on hospital expenditures without providing disincentives for doctor visits this study investigates the

<sup>17</sup> In a separate analysis we investigated whether the precise form of the private insurance cover (with or without deductible) has any effect on the demand for health care. But the exact insurance type does not seem to affect health care demand either.

<sup>18</sup> Our key results on the coefficients of the insurance and moral hazard indicators are robust to distributional assumptions and are confirmed in Poisson, Negbin, and in lognormalP models without cross-equation correlations.



determinants of health care demand. We introduced a new econometric technique to adequately model count data outcomes in a framework of panel data and correlated outcomes. The Poisson estimator with a lognormally distributed random effect for bivariate count processes enhances the efficiency of previously utilized estimators and yields significantly improved estimation results.

In addition to this methodological contribution, this study addressed the consequences of informational asymmetries for health care demand. Whereas adverse selection affects the decision on insurance coverage, moral hazard influences health care demand, conditional on existing insurance coverage. While it is empirically difficult to distinguish adverse selection and moral hazard effects this study attempts to test for the existence of each effect separately.

Within the German health care framework, the most likely suspect for the realization of adverse selection lies in the decision of publicly insured individuals to purchase private supplemental or add-on insurance. Adverse selection takes place if those, who expect to have high demands for health care, are more likely to purchase add-on insurance. Therefore one implication of such an effect would be to find a positive correlation between holding add-on insurance and health care demand. However, this correlation may just as well be an indicator of *ex post* moral hazard behaviour and by itself is insufficient evidence for adverse selection. Therefore we tested whether the decision to purchase add-on insurance, an *ex ante* decision relative to the realized health care demand, is endogenous to subsequent health care demand. If the decision to purchase add-on insurance were exogenous to subsequently realized, but previously expected health care demand, the hypothesis of adverse selection could be rejected. This hypothesis was not rejected in the case of hospital visits which suggests that individuals who expect a large demand for hospital care might have self-selected themselves into add-on insurance. We interpret this finding as supportive of the hypothesis of adverse selection, as theoretically predicted (Jaynes, 1978; Hellwig, 1988).

Our third contribution consists of the analysis of moral hazard effects which may vary depending on the type of health care demanded. Whereas our central indicator of moral hazard behaviour, i.e. whether an individual is privately or publicly insured, does not appear to be predictive of health care demand, we do observe individuals' responses to opportunity costs of health care visits. A striking example is that the self-employed, who in contrast to employees suffer income losses following doctor or hospital visits, visit doctors significantly less often. There is no difference in the demand for hospital visits. The response of females to the presence of children in the household confirms this pattern. While women with children are significantly less likely to visit a doctor, the presence of children in the household does not significantly affect women's demand for hospital visits. These examples suggest a higher elasticity in the demand for outpatient rather than inpatient care, confirming Manning *et al.* (1987).

The policy conclusion of these findings is that if co-payments are introduced as an incentive mechanism to steer health care demand, it is important to impose them on health care services with elastic demands. Our results suggest that the elasticity of demand for hospital care is small and unresponsive to changes in financial and non-financial opportunity costs. If this is the case the introduction of co-payments as an incentive mechanism is pointless for this dimension of health care demand. In addition, to the degree that adverse selection is the determinant of the positive correlation between insurance choice and health care demand, as our results on the add-on insurance seem to suggest for hospital visits, policy instruments addressing moral hazard mechanisms, such as deductibles, will not succeed in affecting health care demand. In contrast to the results on hospital visits, our results on the elasticity of the demand for physician visits confirm the moral hazard effects found in the literature.

## APPENDIX: LOGIT PREDICTION EQUATION: HOLDING ADD-ON INSURANCE

	Males		Females	
	Coefficient	<i>t</i> -value	Coefficient	<i>t</i> -value
<b>Health insurance indicator</b>				
Number of health insurances	0.040**	2.426	0.056**	5.441
<b>Health indicators</b>				
Handicap (0/1)	0.319	0.942	0.180	0.874
Degree of handicap	−0.020**	−1.977	0.003	0.706
<b>Schooling</b>				
Years of schooling	0.810**	3.568	0.471**	4.455
Hauptschul degree	2.537**	2.616	0.160	0.733
Realschul degree	1.489*	1.853	0.840	0.738
Abitur degree	−0.190	−0.338	−0.766	−0.853
Polytechnical degree	—	—	−0.321	−0.365
University degree	−2.113**	−2.647	−0.635	−1.368
<b>Age indicators</b>				
Age 30 to 34	0.422	1.140	0.762**	2.434
Age 35 to 39	0.155	0.363	1.065**	3.439
Age 40 to 44	−0.127	−0.267	0.742**	2.284
Age 45 to 49	−0.374	−0.750	0.835	2.507
Age 50 to 54	−0.400	−0.750	1.224	3.706
Age 55 to 59	−0.990	−1.474	0.827**	2.252
Age > 59	0.926**	2.048	1.058**	0.360
White collar employment (0/1)	0.461*	1.843	−0.325**	−2.020
Married (0/1)	0.356	1.121	−0.119	−0.623
Household net income	0.00019**	2.787	0.00020**	7.269
Children in household (0/1)	−0.274	−0.953	0.185	1.005
Constant	−17.218**	−5.298	−11.645**	−8.582
lnL	−432.58		−955.32	
Number of observations	9,274		11,669	

*Note:* \* and \*\* indicate statistical significance at the 10% and 5% level, respectively. The variable 'polytechnical degree' is excluded from the male model because it perfectly predicts the outcome.

## ACKNOWLEDGEMENTS

We thank two anonymous referees, Jochen Mayer and seminar participants at the University of Munich and at the Deutsches Institut für Wirtschaftsforschung (DIW) in Berlin for helpful comments.

## REFERENCES

- Altman D, Cutler DM, Zeckhauser RJ. 1998. Adverse selection and adverse retention. *American Economic Review* **88**(2): 122–126.
- Cameron AC, Trivedi PK. 1986. Econometric models based on count data: Comparisons and applications of some estimators and tests. *Journal of Applied Econometrics* **1**: 29–53.
- Cameron AC, Trivedi PK. 1998. *Regression Analysis of Count Data*. Cambridge University Press: New York.
- Cameron AC, Trivedi PK, Milne F, Piggot J. 1988. A microeconomic model of the demand for health care and health insurance in Australia. *Review of Economic Studies* **55**: 85–106.
- Chiappori PA, Salanie B. 2000. Testing for asymmetric information in insurance markets. *Journal of Political Economy* **108**: 56–78.
- Chiappori PA, Durand F, Geoffard PY. 1998. Moral hazard and the demand for physician services: first lessons from a French natural experiment. *European Economic Review* **42**: 499–511.
- Coulson NE, Terza JV, Neslusan CA, Stuart BC. 1995. Estimating moral-hazard effects of supplemental medical insurance in the demand for prescription drugs by the elderly. *American Economic Review* **85**: 122–126.
- Ellis RP. 1989. Employee choice of health insurance. *Review of Economics and Statistics* **71**(2): 215–223.
- Geil P, Million A, Rotte R, Zimmermann KF. 1997. Economic incentives and hospitalization in Germany. *Journal of Applied Econometrics* **12**(3): 295–311.
- Grossman M. 1972. On the concept of health capital and the demand for health. *Journal of Political Economy* **80**: 223–255.
- Hausman J. 1978. Specification tests in econometrics. *Econometrica* **46**(6): 1251–1271.
- Hausman J, Hall B, Griliches Z. 1984. Econometric models for count data with an application to the patents-R&D relationship. *Econometrica* **52**: 909–938.
- Hellwig MF. 1988. A note on the specification of interfirm communication in insurance markets with adverse selection. *Journal of Economic Theory* **46**: 154–163.
- Holly AI, Gardiol L, Domenighetti G, Bisig B. 1998. An econometric model of health care utilization and health insurance in Switzerland. *European Economic Review* **42**: 513–522.
- IW (Institut der Deutschen Wirtschaft), various years. *Zahlen zur wirtschaftlichen Entwicklung der Bundesrepublik Deutschland*. Deutscher Instituts-Verlag GmbH: Köln.
- Jaynes GD. 1978. Equilibria in monopolistic competitive insurance markets. *Journal of Economic Theory* **19**: 394–422.
- Long SH, Marquis MS, Rodgers J. 1998. Do people shift their use of health services over time to take advantage of insurance? *Journal of Health Economics* **17**: 105–115.
- Manning WG, Newhouse JP, Duan N, Keeler EB, Leibowitz A, Marquis MS. 1987. Health insurance and the demand for medical care: evidence from a randomized experiment. *American Economic Review* **77**(3): 251–277.
- Marquis MS. 1992. Adverse selection with a multiple choice among health insurance plans: a simulation analysis. *Journal of Health Economics* **11**: 129–151.
- Marquis MS, Phelps CE. 1987. Price elasticity and adverse selection in the demand for supplementary health insurance. *Economic Inquiry* **25**: 299–313.
- Million A. 1998. *Zählmodellen für korrelierte abhängige Daten*. Unpublished dissertation, University of Munich.
- Müller W. 1995. Ausgaben für die Gesundheit 1993. *Wirtschaft und Statistik* **95**(12): 914–922.
- Pohlmeier W, Ulrich V. 1995. An econometric model of the two-part decision-making process in the demand for health care. *Journal of Human Resources* **30**(2): 339–361.
- Preston FW. 1948. The commonness and rarity of species. *Ecology* **29**: 254–283.
- Rothschild M, Stiglitz J. 1976. Equilibrium in competitive insurance markets: an essay on the economics of imperfect information. *Quarterly Journal of Economics* **90**: 629–650.
- STBA (Statistisches Bundesamt). 1997. *Statistisches Jahrbuch für die Bundesrepublik Deutschland*. Metzler-Poeschl Verlag: Wiesbaden.
- van de Ven WPM, van Vliet RCJA. 1995. Consumer information surplus and adverse selection in competitive health insurance markets: an empirical study. *Journal of Health Economics* **14**: 149–169.
- Wagner GG, Burkhauser RV, Behringer F. 1993. The English language public use file of the German socio-economic panel. *Journal of Human Resources* **28**: 429–433.
- Winkelmann R. 2000. *Econometric Analysis of Count Data Models*, 3rd edition. Springer-Verlag: Berlin.