Incentive Effects in the Demand for Health Care: A Bivariate Panel Count Data Estimation

Regina T. Riphahn[1,2,3]

Achim Wambach[1,2]

Andreas Million[1,3]

[1] University of Munich, [2] CEPR, London, [3] IZA, Bonn

April 2000

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.

JEL classification:	I11, C23, C25, D82
Keywords:	Asymmetric information, demand for health services, bivariate panel count data estimation

Correspondence to:
Regina T. Riphahn
University of Munich
Ludwigstr. 28 RG
80539 Munich, Germany
Phone: +49-89-2180 2128
Fax: +49-89-33 63 92
Email: Regina.Riphahn@selapo.vwl.uni-muenchen.de

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

1 Introduction

Since health care expenditures are substantial and rising in most industrialized countries it is important to understand their determinants.[1] 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.[2] 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, behavioral idiosyncracies, such as the propensity to seek health care given a certain ailment, might differ in unobservable but systematic ways between individuals. Second,

[1] The share of German public and private health insurance expenditures in GDP increased from 4.2 percent in 1970, to 6.5 percent in 1990 and to 8.02 percent in 1997 (IW, various years). The share of health related expenditures in the U.S. budget increased from 11.3 to 19.2 percent between 1985 and 1995 (STBA, 1997).

[2] 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.

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. 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 ...".[3] While the results indicate possible response patterns, they do not provide conclusive evidence on actual behaviors. 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).[4] In contrast to the experiment based literature our paper uses representative survey data on actual individual behaviors 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

[3]e.g. Marquis and Phelps (1987), van de Ven and van Vliet (1995); for an exception see Altman et al. (1998)

[4]For a discussion of the problems in establishing empirical evidence for adverse selection, see van de Ven and van Vliet, 1995.

to the analysis of moral hazard, i.e. whether the demand for medical care of insured individuals exceeds that of identical uninsured persons. 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 these services will respond differently to given incentives, e.g. in the form of 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 behavioral 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 three 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 four and five, before part six of the paper reviews the results and section seven 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 6,150 gross monthly earnings in 1997, about $3,500). 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 percent of the population is insured in mandatory insurances and only about 0.1 percent 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, e.g. to join those with the lowest rates. 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.

In addition to pure public or private insurance, add-on insurance can be purchased. People who have public insurance might take up add-on private insurance, which covers extra costs like double rooms in hospital, or reimbursements for copayments on glasses. 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 just as they invest in human capital. There is a stock of health capital, which depreciates over time: $H_{i+1} - H_i = I_i - \delta_i H_i$. δ_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:

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

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: 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

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

with $H = \prod_k e_k^{\alpha_k(s,A,B)}$, which can be interpreted as a health production function. σ 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.

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:

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

where Z is a vector of covariates and D_j are dummy variables for the jth 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 η_{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 endogenity 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 explicitely, 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 inhowfar 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 provides 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[5], from the West German subsample, and of German nationality. After dropping observations with missing values on key variables the sample contained 3,691 male and 3,689 female individuals, which make up a sample of 14,243 male and 13,794 female person-year observations.[6]

[5] By this age restriction we exclude retirees and individuals who are still covered by their parents' health insurance.

[6] Observations of women who gave birth within the period of six months before and after a given calendar year were omitted from the sample.

Our dependent variables are the number of visits to a doctor within the last quarter prior to the survey, and the number of hospital visits 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.[7] Both dependent variables are integer valued. Descriptive statistics are presented in Table 1 by gender. Hospital visits are highly skewed at zero, and the average number of physician visits amounts 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 percent 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 percent in the male and 9 percent 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 percent of the publicly insured subsamples.

For descriptive statistics on these and other explanatory variables see Table 2. Clear differences between the male and female subsamples are found with respect to

[7] 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.

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 percent of the male sample is employed compared to 49 percent of females, and among employed men the fractions who are self employed, civil servants, or blue collar workers are higher.

Table 3 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 3 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 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 λ_i, the so-called intensity. Covariates are introduced using an exponential parametrization of λ_i such that the expected value of y_i is a function of the covariates x_i and coefficients β:

$$
\begin{aligned}
y_i \sim Po(\lambda_i) \quad : \quad p(y_i = a) \; &= \; \frac{e^{-\lambda_i} \lambda_i^a}{a!} && (3) \\
E(y_i) \; = \; V(y_i) \; = \; \lambda_i \; &= \; e^{\beta' x_i} && (4)
\end{aligned}
$$

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 λ_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:

$$
\begin{aligned}
p(y_i = a | \lambda_i) \; &= \; \frac{e^{-\lambda_i} \lambda_i^a}{a!} && (5) \\
\ln(\lambda_i) \; &\sim \; N(\beta' x_i, \, \sigma) && (6)
\end{aligned}
$$

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

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

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 lable 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:

$$\ln(\lambda_i) = \beta' x_i + \epsilon_i, \tag{8}$$

where ϵ_i is a normally distributed error term. The right hand side of equation (8) is familiar from linear models or from nonlinear 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 ϵ_i. Also, new models can easily be built by customizing the correlation between different errors ϵ_i and ϵ_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 ϵ_i is pointed out by Winkelmann (1997): 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.

$$\begin{aligned}
y_{itg} &\sim Po(\lambda_{itg})\ ,\quad g = 1, 2 && (9)\\
\ln(\lambda_{it1}) &= \beta_1' x_{it1} + u_{i1} + \epsilon_{it1} && (10)\\
\ln(\lambda_{it2}) &= \beta_2' x_{it2} + u_{i2} + \epsilon_{it2} && (11)\\
(\epsilon_{it1}, \epsilon_{it2}) &\sim N_2(0, 0, \sigma^2_{\epsilon 1}, \sigma^2_{\epsilon 2}, \rho) && (12)\\
u_{i1} &\sim N(0,\ \sigma^2_{u1}) && (13)\\
u_{i2} &\sim N(0,\ \sigma^2_{u2}) && (14)\\
E[\epsilon_{itg} u_{jh}] &= 0 \quad \forall\, i, t, g, j, h && (15)\\
E[\epsilon_{itg} \epsilon_{jsh}] &= 0 \quad \text{if } t \neq s \vee i \neq j \vee g \neq h && (16)\\
E[u_{ig} u_{jh}] &= 0 \quad \text{if } i \neq j \vee g \neq h && (17)
\end{aligned}$$

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 ϵ follow a bivariate normal distribution, while the outcome specific error terms u represent individual specific unobserved heterogeneity, which is constant over time. Statements (12) through (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: 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 e.g. 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, however, our data does not provide the necessary information. The error terms ϵ_{itg} control for shocks which in any period may affect both outcomes. Such correlated period shocks can result e.g. from flue 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 ϵ_{it1} and ϵ_{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, e.g. for cases where 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.

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 ϵ 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 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 β_c is a semi-elasticity. An increase of $\mathrm{d}x_c = 0.01$ changes the expectation of y by β_c percent:

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

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

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

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

6 Results

Model Performance - To judge the merit of the bivariate panel approach against simpler estimation procedures we estimated both models: Table 4 lists the results for the univariately estimated models without random effects. Comparing these results with those of the bivariate panel estimation in Table 5 yields various differences.[8] 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 5 and insignificant in Table 4. 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:

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

[8] The specifications in Tables 4 and 5 were estimated controlling for a set of year fixed effects. These coefficients are omitted in the tabulations to save space.

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

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 σ_u and ρ, which is found for the individual parameters already in Table 5.[9] Comparing estimation results for σ (Table 4) and σ_ϵ (Table 5) shows 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.[10]

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 5 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

[9]The critical value for the likelihood ratio test at the 1 percent significance level with two degrees of freedom is 10.6.

[10]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 one percent 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.

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 0.55 and for women by 0.22 percent, which are large effects. Interestingly, only the coefficient for the male sample is statistically significant, confirming the differential univariate effects already observed in Table 3. Therefore this first evidence appears to corroborates 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.[11] Least squares regressions of doctor and hospital visits

[11] 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. For

then controlled for the variables as in Table 5 plus the predicted values of add-on insurance, omitting only the indicator variable for public insurance. Table 6 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 ten percent 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.[12]

This evidence compares well with the results presented by Holly et al. (1998), who strongly reject the exogeneity of the insurance purchase to subsequent health care utilization, using Swiss cross section data. Also, Ellis (1989) using a very different approach confirmed that those individuals with high health care expenditures in the past, choose insurance plans with extensive coverage.

Moral Hazard - Turning now to the incentives affecting demand for health care, moral hazard driven behavior 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,[13] however the hypothesis could not be confirmed in panel data models for Germany so far.[14] Therefore we generalize the test for

---

two reasons we focus on the former group of individuals who are mandatorily in the public insurance: First, this provides a homogenous sample of individuals with a comparable decision framework, and second, if we were to consider the decision of the second group additionally, it would have to be modelled differently and were unlikely to yield reliable results due to the small number of observations.

[12]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.

[13]See e.g. Long et al. (1998), Coulson et al. (1995), or for a classic Manning et al. (1987).

[14]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

moral hazard behavior applied by Geil et al. (1997) by considering different outcomes of health care demand. Our estimation results (see Table 5), 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.[15] 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) percent 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 young 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

a specialized physician.

[15]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.

to the increased opportunity cost of health care demand brought about by the presence of children.

Other Findings - Turning now to the interpretation of the results in Table 5, 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 3. 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 labor 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 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 behavior 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). This establishes a second important contribution to the literature, as so far very little non-experimental evidence is available on the existence of such effects.

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 behavior, 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 conclusions of these findings are clear: 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. If, as our findings suggest, the elasticity of demand for hospital care is small and unresponsive to changes in financial and non-financial opportunity costs, then the introduction of co-payments as an incentive mechanism is pointless. In addition, to the degree that adverse selection is the determinant of the positive correlation between insurance choice and health care demand, policy instruments addressing moral hazard mechanisms, such as deductibles, will not succeed in affecting health care demand.

References

Altman, D., D.M. Cutler, and R.J. Zeckhauser, 1998: Adverse Selection and Adverse Retention, *American Economic Review* 88(2), 122-126.

Cameron, A. C., Trivedi, P. K., 1986: Econometric models based on count data: Comparisons and applications of some estimators and tests, *Journal of Applied Econometrics* 1, 29-53.

Cameron, A. C., and P.K. Trivedi, 1998: *Regression Analysis of Count Data*. New York: Cambridge University Press

Cameron, A. Colin, P.K. Trivedi, F. Milne, and J. Piggot, 1988: A Microeconomic Model of the Demand for Health Care and Health Insurance in Australia, *Review of Economic Studies* 55, 85-106.

Chiappori, P.A., F. Durand, and P.Y. Geoffard, 1998: Moral Hazard and the Demand for Physician Services: First Lessons from a French Natural Experiment, *European Economic Review* 42, 499-511.

Coulson, N.E., J.V. Terza, C.A. Neslusan, and B.C. Stuart, 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, R.P., 1989: Employee Choice of Health Insurance, *Review of Economics and Statistics* 71(2), 215-223.

Geil, P., A. Million, R. Rotte and K.F. Zimmermann, 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., B. Hall, and Z. Griliches, 1984: Econometric models for count data with an application to the patents-R&D relationship, *Econometrica* 52, 909-938.

Hellwig, M.F., 1988: A Note on the Specification of Interfirm Communication in Insurance Markets with Adverse Selection, *Journal of Economic Theory* 46, 154-163.

Holly, A., L. Gardiol, G. Domenighetti, and B. Bisig, 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*, Köln: Deutscher Instituts-Verlag GmbH.

Jaynes, G.D., 1978: Equilibria in Monopolistic Competitive Insurance Markets, *Journal of Economic Theory* 19, 394-422.

Long, S.H., M.S. Marquis, and J. Rodgers, 1998: Do People Shift Their Use of Health Services Over Time to Take Advantage of Insurance, *Journal of Health Economics* 17, 105-115.

Manning, W.G., J.P. Newhouse, N. Duan, E.B. Keeler, A. Leibowitz, and M.S. Marquis, 1987: Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment, *American Economic Review* 77(3), 251-277.

Marquis, M.S., 1992: Adverse Selection with a Multiple Choice Among Health Insurance Plans: A Simulation Analysis, *Journal of Health Economics* 11, 129-151.

Marquis, M.S. and C.E. Phelps, 1987: Price Elasticity and Adverse Selection in the Demand for Supplementary Health Insurance, *Economic Inquiry* 25, 299-313.

Million, A., 1998: *Zähldatenmodelle 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. and V. Ulrich, 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, F. W., 1948: The Commonness and Rarity of Species, *Ecology* 29, 254-283.

Rothschild, M. and J. Stiglitz, 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, W.P.M.M. and R.C.J.A. van Vliet, 1995: Consumer Information Surplus and Adverse Selection in Competitive Health Insurance Markets: An Empirical Study, *Journal of Health Economics* 14, 149-169.

Wagner, G. G., R. V. Burkhauser, and F. Behringer, 1993: The English Language Public Use File of the German Socio-Economic Panel, *Journal of Human Resources* 28, 429-433.

Winkelmann, R. 1997: *Count Data Models: Econometric Theory and an Application to Labor Mobility*, second edition; Berlin, Heidelberg: Springer-Verlag.

Table 1: Dependent Variables - Descriptive Statistics

	Hospital Visits		Doctor Visits	
	Male	Female	Male	Female
Value	(Share of Total Observations, in Percent)			
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
N	14,243	13,083	14,243	13,083

Source: German Socioeconomic Panel (1984-1995)

Table 2: List of variables

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

[a] means, standard deviations in parenthesis

Source: German Socioeconomic Panel (1984-1995)

Table 3: 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				
no	2.50	3.63	0.176	0.187
yes	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)

Table 4: LogNormalP

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

t-values are given in parentheses.
Two-sided test: * 10% significance, ** 5% significance

Table 5: bivariate PanelLogNormalP

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

t-values are given in parentheses.
Two-sided test: * 10% significance, ** 5% significance

Table 6: 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 5, only public ins is excluded.