

Insurance Coverage and the Heterogeneity of Health and Drug Spending in the United States

Marin Gemmill^a, Joan Costa-Font^{a,b} and Panos Kanavos^a

^aLSE Health, London School of Economics, Cowdray House, Houghton Street, WC2A 2AE, London, U.K
E-mails: J.Costa-Font@lse.ac.uk, M.C.Gemmill@lse.ac.uk, P.G.Kanavos@lse.ac.uk

^bDepartament de Teoria Econòmica, Universitat de Barcelona, Catalonia, Spain.

The rise in health expenditure in the U.S. has generated interest in the determinants of health expenditure at the micro-level; however, individuals may exhibit differential behaviour across different types of health care. In addition, public and private insurance may have an heterogeneous effect on expenditure. In this paper, we examine the determinants of health expenditure with a multivariate regression model along with controls for insurance choice and unobserved health care utilization. We find age-related effects and evidence of moral hazard related to private insurance, while the primary effect of income on expenditure appears to be through the purchase of insurance. The implications of the study are that: (i) policymakers should be less concerned about the effect of ageing on health expenditure; (ii) drug spending may not be related to the expansion of public insurance coverage; and (iii) income may have a negative impact on most elements of health spending.

The Geneva Papers (2006) 31, 669–691. doi:10.1057/palgrave.gpp.2510098

Keywords: health insurance; health expenditure; multivariate regression; sample selection; endogeneity

Introduction

Health expenditure in the United States (U.S.) stands at approximately 15 per cent of GDP, and the expenditure growth rate has accelerated from a low of 1.9 per cent in 1995 to a peak of 9.3 per cent between 2001 and 2002.¹ However, there is significant heterogeneity in the patterns of expenditure growth; for instance, although pharmaceutical expenditure comprised approximately 10 per cent of total national health spending in 2004, this expenditure component contributed 14.7 per cent of total health care spending growth from 1994 to 2004.² While it has been established that insurance coverage is an important determinant of health expenditure growth,³ whether the effects of insurance are heterogeneous across different types of health care is an empirical question that remains unanswered. If insurance induces consumers to substitute certain forms of care (e.g., inpatient care) for alternative treatments (e.g., pharmaceuticals), the impact of health insurance may vary by the type of health expenditure. There may also be heterogeneous effects

¹ OECD (2005).

² KFF (2006b).

³ Newhouse (1993).

associated with other demographic and socio-economic variables. For example, if age plays a role in explaining the heterogeneity of health and pharmaceutical expenditure, then while inpatient expenditure may vary significantly with proximity to death, pharmaceutical expenditure might not. Because of this possible heterogeneity, it is unclear whether previous empirical results hold if several forms of expenditure are jointly determined, suggesting that an alternative course of action, which captures possible correlations in the error terms of different equations, should be considered.

Insurance coverage may not be independent of health expenditure given that the decision to obtain insurance is influenced by the individual's expectations of how much medical care he needs to consume and given that insurance coverage may affect the prices of health care.⁴ In other words, the insurance decision is potentially endogenous. Although researchers have tested for the endogeneity of insurance coverage in the context of drug expenditure,⁵ the issue is still unresolved. Furthermore, we frequently observe health expenditure for those individuals who need health care and choose to purchase it at the point of need, but failing to correct for this effect may lead to biased estimates.⁶

An analysis of various health expenditure components is timely as the current U.S. administration is pursuing a number of reforms to reduce health expenditure. Proposals include changing how states run the Medicaid programme, extending private insurance coverage through increased tax incentives, allowing insurance companies to operate across state lines, permitting small businesses to pool and purchase health insurance, and increasing Medicare premiums for high-income recipients.⁷ Whether these proposals will offer much relief from health expenditure growth without hurting the most vulnerable populations is an open question. For example, if expenditure increases are driven by private insurance rather than public insurance, then proposals to increase private insurance coverage through tax credits could have the opposite effect of that intended. Moreover, targeting premium increases and coverage reductions at vulnerable populations (e.g., the elderly or the poor) could be harmful if there are barriers in access to care for these groups or if utilization is primarily driven by need rather than incentive inefficiencies like moral hazard. A better understanding of how public and private insurance influences expenditure and how various demographic, socio-economic, and health variables contribute to health expenditure would aid policymakers in determining the likely impact and success of specific proposals.

The main objective of this paper is thus to examine the factors that determine health expenditure, mainly insurance coverage and other socio-economic, demographic, and health variables, and to determine if these factors have differential effects on the main components of health expenditure, notably inpatient, outpatient, pharmaceutical, emergency care, and office visits. The paper contributes to the existing

⁴ Frank (2001).

⁵ Lillard *et al.* (1999).

⁶ Vella (1998).

⁷ The White House (2006); KFF (2006a).

literature by employing an alternative methodology (multivariate regression analysis) to account for the heterogeneity of health expenditure and by further controlling for the endogeneity of insurance and sample selection. It draws upon the 1996–2003 Medical Expenditure Panel Survey, which contains a rich amount of private and public insurance expenditure information. Based on the main results from the analysis, the paper also seeks to contribute to the current policy debate in the U.S.

The following section briefly outlines the institutional background of health insurance coverage in the U.S. and offers a general overview of the empirical evidence relating to insurance and other predictors of health expenditure. The next section details the econometric model and discusses the endogeneity and sample selection corrections along with the intuition behind the multivariate regression. The following section discusses the data and presents the descriptive results, whereas the penultimate section reports the results of the econometric analysis. Finally, the last section concludes by discussing the policy implications.

Institutional background and evidence

Insurance coverage in the U.S. is fragmented with individuals receiving coverage from various private and public third-party payers and 16 per cent of the population having no insurance coverage.⁸ Of the public third-party payers, Medicare mainly covers the elderly and certain other groups.⁹ Hospital coverage is automatically extended to all beneficiaries, while outpatient and physician services and some medical services are available under voluntary Medicare coverage. Until 2006 most outpatient prescription drugs were not covered by the programme. The state-run Medicaid programme is intended to cover low-income and disabled Americans and their families, although each state establishes its own eligibility requirements.¹⁰ States have considerable freedom in determining eligibility for coverage, although there are certain groups that states must cover in order to receive Federal funds.

Private insurance coverage is mainly in the form of employer-sponsored insurance, as there are tax incentives for employers that offer health insurance.¹¹ Virtually all covered workers in employer-based plans (98 per cent) had a prescription drug benefit in 2005. While approximately 60 per cent of the population had employer-sponsored insurance in 2004, only around 9 per cent of the population had non-group (individual) insurance,¹² which is an alternative form of coverage for those without access to employer-sponsored coverage. Only 80 per cent of enrollees with non-group health insurance had prescription drug coverage in 2000.¹³

⁸ U.S. Census Bureau (2005).

⁹ CMS (2006).

¹⁰ *Ibid.*

¹¹ KFF/HRET (2005).

¹² US Census Bureau (2005).

¹³ Gabel *et al.* (2002).

The relationship between health insurance and expenditure

A significant body of literature has researched the effect of insurance on medical care expenditure. The RAND experiment was a key study, and a number of studies¹⁴ have established that more generous insurance led to increased health care expenditure among the non-elderly using this data. Using alternative datasets, other researchers¹⁵ have found that insurance coverage led to higher medical expenditure in the elderly. A number of other covariates have been found to increase health expenditure, including having higher income,¹⁶ more education,¹⁷ and being in poor health,¹⁸ although being female led to lower health expenditure.¹⁹

Some research has focused exclusively on the determinants of pharmaceutical expenditure, and the general finding is that insurance coverage increases total pharmaceutical spending. One finding is that elderly persons with private drug coverage had higher drug expenditure than those only covered by Medicare.²⁰ In addition, the type of insurance coverage was a significant predictor of expenditure: conditional on having at least one prescription, elderly Medicaid recipients had lower drug expenditure than individuals with no supplemental coverage.²¹

In addition, researchers have found that elderly individuals with more limitations to activities of daily living (ADLs) and persons diagnosed with angina, congestive heart failure, hypertension, diabetes, or arthritis had higher prescription drug expenditure.²² Furthermore, the more individuals in a household who claimed to be in poor health, the higher the pharmaceutical spending for that household.²³ Few other demographic variables have been found to impact expenditure, although expenditure was higher for elderly persons living in urban areas.²⁴ Other research has established that adult women spent twice as much on prescription medicines as men²⁵ and increases in the household size led to lower household prescription drug expenditure.²⁶ The general findings from the literature regarding pharmaceutical expenditure are reported in the Appendix.

The model

There are several important theoretical and empirical considerations related to the determinants of pharmaceutical and other health-care expenditure, which include

¹⁴ Manning *et al.* (1981); Newhouse *et al.* (1981); Duan *et al.* (1983); Newhouse (1993).

¹⁵ Christensen *et al.* (1987); Cartwright *et al.* (1992).

¹⁶ Manning *et al.* (1981).

¹⁷ Cartwright *et al.* (1992).

¹⁸ *Ibid.*

¹⁹ *Ibid.*

²⁰ Long (1994); Federman *et al.* (2001); Yang *et al.* (2004).

²¹ Yang *et al.* (2004).

²² Lillard *et al.* (1999).

²³ Street *et al.* (1999).

²⁴ Lillard *et al.* (1999).

²⁵ Liebowitz *et al.* (1985).

²⁶ Smith (1993); Street *et al.* (1999)

the endogeneity of the insurance decision, the unobserved consumption for individuals who do not consume any medical care in a given year, and the inter-dependency of different sources of expenditure. The methodological approach proposed in this study relies on three main corrections. First, we correct for the endogeneity of insurance coverage by estimating a multivariate insurance decision, separated into the possible insurance coverage decisions within the U.S. healthcare context: Medicare only, other public insurance (includes individuals with Medicare and another form of public coverage, Medicaid, TRICARE, VA health insurance, or some other form of public coverage), private insurance (includes individuals with Medicare and private coverage or private coverage only), and no insurance. We then insert the estimated predictions into the expenditure equations after bootstrapping the standard errors. The rationale behind the correction is that because public insurance is intended for the most vulnerable populations, those with private insurance tend to be healthier, to have higher incomes and better education, and to be in full-time employment.²⁷ More specifically, insurance coverage may influence the consumption of medical care, but the amount of medical care that an individual *expects* to consume may influence his choice of insurance coverage. Moreover, we anticipate that individuals with private insurance would be less in need of medical care than those with public insurance, and indications of higher consumption among those with private insurance may be more reflective of moral hazard than need.

Second, given that health expenditure is significantly skewed due to the infrequency of the insurance purchase, the probability of any expenditure is estimated separately (the “selection equation”), and the sample selection correction (Mills lambda) is inserted into the expenditure equation to correct for selection bias.²⁸ This requires an identification variable that explains the generation of expenditure but is not related to the participation equation. The rationale behind the sample selection correction is that the individual makes marginal cost–marginal benefit comparisons when faced with the decision on whether to seek medical care. Because of the inability to observe whether a value of zero for expenditure is due to the individual never needing medical care (not a true zero) or the individual being prescribed treatment and choosing not to purchase it (a true zero), sample selection techniques are employed.

An alternative to the Heckman methodology is the two-part model,²⁹ which was created to model the demand for health care and is motivated by the conditional mean independence assumption:

$$E(y_i | y_i > 0, x_{2i} = x_{2i}\beta_2) \tag{1}$$

There has been extensive debate regarding the choice between the two-part model and the sample selection model.³⁰ One argument is that two-part models are more

²⁷ Feldman *et al.* (1989); Fronstin *et al.* (1997).

²⁸ Heckman (1979).

²⁹ Duan *et al.* (1983).

³⁰ Jones (2001).

appropriate for sequential decisions, although Maddala³¹ argues that even when decisions are sequential, the decisions will be correlated if there are omitted variables common to both. Leung and Yu³² used Monte Carlo simulations to compare these two differing methods and determined that the choice of model depends on the empirical context. Specifically, the sample selection model performs poorly when there is collinearity, which can arise in certain contexts: when there is a large degree of censoring, when there are few exclusion restrictions, when there is little variability between regressors, or when there are weak instruments.³³ In the absence of collinearity, which can be verified using the condition number, the *t*-test on the inverse Mills ratio is an indicator of which of the two specifications is more appropriate.

Third, because we estimate the determinants of different expenditure sources simultaneously, the expenditure equation is a multivariate regression model rather than a standard ordinary least squares (OLS) model. Multivariate regression models are a more general form of an empirical equation that allow for more than one endogenous variable. The advantage is that the several sources of expenditure can be disaggregated, permitting us to jointly estimate the model and control for possible unobserved heterogeneity. These models are typically implemented when there are two or more strongly correlated (paired) dependent variables. This is the case with health care decision-making where, for example, an individual's drug consumption depends on previously obtaining a prescription from the physician or hospital.

The insurance equation

We first proceed to estimate the determinants of insurance coverage using a multinomial logistic model where the individual chooses between Medicare only, other public insurance, private insurance, or no insurance. Let I_{ik} represent a vector of $k=4$ insurance alternatives ($I_i^{MO}, I_i^{OP}, I_i^{PR}, I_i^{NI}$), where MO =Medicare only, OP =other public, PR =private, and NI =no insurance. Let w_i denote a set of insurance determinants where it is possible to obtain the probability of each outcome or insurance option conditioned on a set of covariates $p(I_i^{MO}, I_i^{OP}, I_i^{PR}, I_i^{NI}/w_i)$ such that the sum adds to unity. Accordingly:

$$p(I_i^k = k) = \exp(\gamma w_i) / \left[1 + \sum_k \exp(\gamma w_i) \right] \quad (2)$$

Following Wooldridge,³⁴ the fitted probabilities can be used for prediction purposes (\hat{I}_i^k).

³¹ Maddala (1985).

³² Leung and Yu (1996).

³³ *Ibid.*

³⁴ Wooldridge (2002).

The participation equation

After estimating the insurance decision, the next step is to correct for possible sample selection bias. We designate health expenditure by $y_i^* = x_i\beta + \mu_i$, where $j=1, \dots, 5$ refers to each type of expenditure measured as a latent variable (y^*) and i represents different individuals across the 8 years of the sample. The observed probability of non-negative health expenditure is:

$$d_i^* = \alpha Z_i + \varphi \hat{I}_i^k + \varepsilon_i, \text{ where } d_i = 1 \text{ if } d_i^* > 0 \text{ and } d_i = 0 \text{ otherwise} \quad (3)$$

where Z_i are vectors of exogenous variables measuring the probability of any expenditure, ε_i refers to the normally distributed error terms with zero mean, and $E(\mu_i/\varepsilon_i) \neq 0$. Hence, we estimate the Mills lambda ($\hat{\lambda}$), which is used to correct for possible sample selection in the expenditure equation via the traditional two-step procedure.³⁵ Finally, we define $y_i = d_i^* y_i^*$, which captures the observed health expenditure.

Expenditure equation and multivariate regression

Health care expenditure is affected by the presence of significant unobserved heterogeneity that results from the fact that expenditure on one source (e.g., drugs) is jointly formed with expenditure from another source (e.g., medical visits). Based on this observation, one possibility is to estimate a model in which multiple response variables are taken as dependent variables. The multivariate regression estimates the same coefficients and standard errors as that of separate OLS models, and given the joint estimation of the relevant equation, the multivariate regression also estimates the between-equation covariances, allowing us to test coefficients across equations. As a result, the only difference from the OLS regression is that the covariance matrix resulting from the dependent variable is a matrix of variables and not just a single vector. Multivariate regression models can be envisaged as a reduced form of a simultaneous equation model.³⁶ Define an n -dimensional ($n=5$) vector of expenditure is source variables y_{ni} by:

$$y_{ni} = \beta x_i + \delta \hat{I}_i^k + \gamma \hat{\lambda}_i + \mu_i \quad (4)$$

where x_i is an n^*k -dimensional vector of expenditure determinants for each individual, λ is the n -dimensional sample selection correction, and μ_i is an n -dimensional random term made up of unknown parameters $N(0, \Sigma)$.

The data and descriptive statistics

The data

The study employs data from the 1996–2003 Medical Expenditure Panel Survey (MEPS), a study co-sponsored by the Agency for Healthcare Research and Quality

³⁵ Vella (1998).

³⁶ Amemiya (1974).

(AHRQ) and the National Center for Health Statistics (NCHS). We also gathered data from the 2006 *Statistical Abstract of the United States* to create certain instrumental variables (IVs). The MEPS survey contains a nationally representative sample of the U.S. civilian, non-institutionalized population with over-sampling of Hispanics and blacks.³⁷ Each year, half of the households are replaced by a new sample of households, creating a longitudinal database. The dataset contains a number of socio-economic and demographic variables along with information on health status, expenditure on various categories of medical care, insurance coverage, and access to care.

Although MEPS contains household data, the data is also available at the individual level. As coverage for medical care is sometimes individual rather than group, which is the case with Medicare, Medigap, and other types of private insurance where the policyholder does not extend coverage to dependents, the unit of analysis in this research is the individual. The health expenditure variables were calculated by combining self-reported data on health care use with cost information reported by health care providers. Specifically, respondents were first asked to identify all medical visits and prescription drug purchases along with the medical practitioners associated with this care and the pharmacies where prescriptions were obtained. Then, the survey administrators contacted each relevant medical practitioner and pharmacy to obtain payment information.

As the panel was relatively short at only two years, we chose the most recent observation for each individual in the sample and pooled the cross-sectional data from 1996–2003. We also excluded respondents under the age of 19 years since a child's guardian typically chooses his health insurance coverage and because children are usually covered under a parent's health insurance policy or through public insurance. After excluding observations with missing values and individuals under the age of 19 years, our sample consisted of $N=91,039$ persons.

The MEPS database has a number of advantages over other U.S. databases that contain health care information, specifically the Panel Study of Income Dynamics (PSID), the National Ambulatory Medical Care Survey (NAMCS), and the Medicare Current Beneficiary Survey (MCBS). The PSID obtains information on medical expenditure by asking the respondent to recall his medical outlays during the past two years, which may lead to rough estimates of medical expenditure that are subject to considerable biases. Another shortcoming of the PSID survey is that expenditure is not separated into different components. Meanwhile, the NAMCS lacks information on prescription drug expenditure along with expenditure on other forms of health care and socio-economic variables. Finally, the MCBS only consists of individuals covered under the Medicare programme, mainly the elderly and some of the disabled. As disease patterns often commence before the start of Medicare eligibility, and this onset of chronic illness has an important impact on health expenditure, we were interested in including the non-elderly in our sample.

³⁷ AHRQ (2004).

Explanatory variables

The impact of insurance on health expenditure can occur via several pathways. First, there is the moral hazard effect of insurance where individuals who face a lower marginal cost of care under insurance consume more health care than if they faced the full price of care.³⁸ Second, patients may have little incentive to consume lower-priced treatments as they do not face the full costs of these treatments; however, third-party payers have mechanisms such as gatekeeping, prior authorization, and tiered co-payments to partially counteract this effect. Third, insurance might influence the agency relationship,³⁹ where physicians and hospitals may have limited incentives to curtail unnecessary medical care. Accordingly, insurance companies have implemented managed care programmes that often pay physicians a capitated amount per patient rather than a fee-for-service amount per good provided or service rendered, or maintain fee-for-service reimbursement, subject to a volume-based adjustment of the relevant fees. Fourth, insurance influences the price of care; specifically, insurance companies, Pharmacy Benefit Managers, and public programmes tend to exert significant negotiating power on providers, and the cost of medical care is lower for third-party payers than for uninsured patients.⁴⁰ Although we would expect a positive effect of insurance on the probability of obtaining medical care, the overall effect of insurance on expenditure is unclear as it depends on whether the volume effect outweighs the price effect of insurance.

On the basis of the existing literature, income is expected to increase the probability of private insurance and negatively influence the probability of non-insurance, although income should have a negative effect on public insurance (excluding individuals with Medicare only). Age and gender are likely to determine the type of expenditure and insurance because of time horizon effects and age- and gender-related risk attitudes. We also expect poor health to increase the probability of insurance and health-care use.

Table 1 provides aggregate evidence suggesting three specific features. First, across all age groups and all types of insurance, those who have insurance coverage spend more on average than those who are uninsured. Second, expenditure appears to rise with age, although the effect is combined with that of insurance coverage. Finally, when we restrict our sample to those who have consumed any health care, expenditure is higher than for the consumption group.

The dependent variables are the different types of expenditure, including prescription drugs, outpatient care, inpatient stays, emergency care, and office-based visits (all measured in logs). We find that the two largest components of health expenditure are drugs and office-based visits, which is expected as both are often consumed together. The independent variables were chosen according to their theoretical soundness, although we excluded other variables due to significant multicollinearity following the standard tests (condition index and variance inflation

³⁸ Pauly (1974).

³⁹ Hurd and McGarry (1997).

⁴⁰ Frank (2001); Kanavos and Gemmill (2004).

Table 1 Health and pharmaceutical expenditures by age (1996–2003)

Source of payment	Total health		Total health (health exp > 0)		Total pharmaceutical		Total pharmaceutical (health exp > 0)	
	<65	≥65	<65	≥65	<65	≥65	<65	≥65
Other public insurance	\$2,865	\$4,949	\$3,461	\$5,242	\$702	\$1,385	\$848	\$1,467
Medicare only	\$5,043	\$3,839	\$5,831	\$4,169	\$1,363	\$940	\$1,576	\$1,021
Private Insurance	\$1,452	\$4,445	\$1,887	\$4,638	\$335	\$1,032	\$435	\$1,077
No insurance	\$532	\$896	\$1,146	\$1,408	\$136	\$312	\$292	\$491

factor). (More information regarding the specific variables chosen for the model is available in the Appendix.)

The model also includes gender, which explains expenditure jointly with age effects. Besides demographic variables, we have included socio-economic variables, such as marital status, that determine health production and could be associated with the costs of health information. We also consider income per person to determine the individual's budget constraint for health and other goods. Income was defined in terms of poverty categories as employed by Medicaid. We also controlled for health status as insurance companies might engage in cherry picking and upwardly adjust the premium for those who are more likely to be sick, while poorer health status is related to higher expenditure. Another covariate is insurance which is split into Medicare only, other public coverage, private insurance, and none. Finally, the model includes time controls to capture any inter-temporal variability.

Preliminary evidence

Figure 1 exhibits health expenditure patterns by age groups. Consistent with Zweifel *et al.*,⁴¹ inpatient expenditure increases steeply after the age of 60 years, suggesting that proximity to death matters; however, other forms of expenditure increase more consistently with age.

Figure 2 reveals that despite some changes in the distribution of insurance, the share of insurance coverage from 1996 to 2003 has not significantly changed over the period. However, Figure 3 suggests that both public and private insurance expenditure have increased significantly, implying that expenditure increases are related more to factors such as moral hazard than to an expansion of insurance coverage.

Results

In discussing the determinants of insurance coverage, drug consumption, and health expenditure, our focus is on the performance of the model and the

⁴¹ Zweifel *et al.* (1999, 2004).

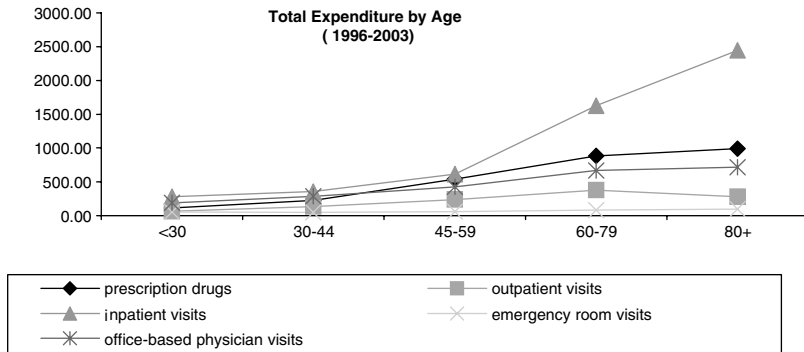


Figure 1. Total expenditure by age (1996–2003) in 1996 U.S.\$

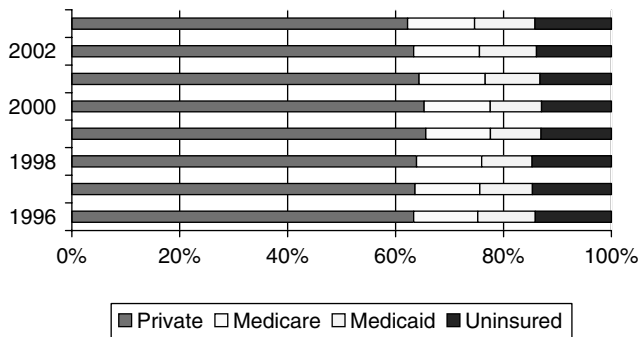


Figure 2. Per cent of U.S. individuals by source of cover. *Note:* individuals can have more than one type of coverage such that the percentages add up to more than 100. *Source:* U.S. Department of Commerce, 2006.

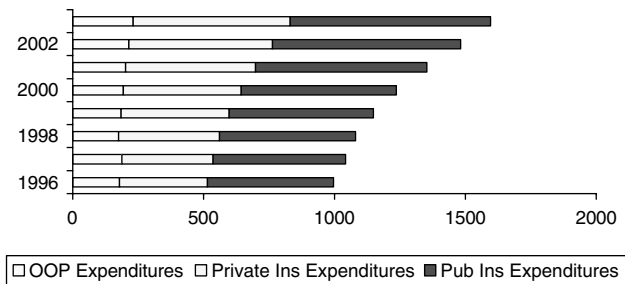


Figure 3. Total value of different sources of insurance coverage in 1996 U.S.\$
Source: U.S. Department of Commerce, 2006.

factors which contribute to the various types of health expenditure. The first stage of the model predicts insurance coverage. Because we hypothesized that insurance was endogenous to the expenditure model, we employed the Smith⁴² test for the weak exogeneity of insurance, which accounts for the censored nature of the data, and we rejected the hypothesis of exogeneity for each expenditure source. We also tested the exogeneity of insurance for each type of expenditure using the Hausman test, and again, we rejected the null hypothesis that insurance was exogenous. Therefore, we proceeded with a two-stage strategy by including the predictions from the insurance equation in the expenditure equation with adequate identification variables and bootstrapped standard errors. When making statistical inferences from finite sample sizes where there is an asymptotic distribution, bootstrapping is used to improve the reliability of standard error estimates.⁴³ We employed no more than 1,000 replications. Also, given the significant share of zeros in the data, we corrected for sample selection by restricting the sample to individuals with positive health expenditure and including the inverse Mills ratios from the predicted probability of positive expenditure in the main expenditure equations.

One of the computational difficulties in our empirical strategy is the selection of adequate instruments. Instrumental variable (IV) models overcome measurement error of several types, including simultaneity and omitted variables, and produce asymptotically unbiased parameter estimates by using an instrument that is not correlated with the measurement or equation error but correlated with the correctly measured variable.⁴⁴ Importantly, the instruments employed in this study are measured at the regional level which could produce a clustering problem and biased standard errors.⁴⁵ To correct for this possibility, the standard errors in both the insurance and probability of expenditure equations were made robust using the White⁴⁶ correction. However, no attempts were made to draw any inferences from the coefficients on the regional instruments as these variables were used exclusively for instrumentation purposes.

The insurance equation contains three instruments, which are highly correlated with the insurance decision, are conceptually relevant, and exhibit a negligible correlation with the expenditure generation process. The first instrument is the percentage of firms in a region that offer health insurance to their employees, which is a measure of price competition in the area. A second instrument is the regional health insurance contribution, namely the average regional family premium per enrolled employee in private companies. This variable likely influences the uptake of different insurance types in the region. The third instrument is the regional age-adjusted mortality rate as there is no evidence of correlation between this instrument and health expenditure, an observation that is consistent with previous work.⁴⁷ We

⁴² Smith (1987).

⁴³ Mills and Zandvakili (1997).

⁴⁴ Angrist and Krueger (2001).

⁴⁵ Moulton (1986).

⁴⁶ White (1980).

⁴⁷ Skinner and Wennberg (2000).

Table 2 Probability of insurance coverage (multinomial logit)

<i>Variable</i>	<i>Other public coverage</i>		<i>Medicare only</i>		<i>Private coverage</i>	
	<i>Coeff</i>	<i>z-value</i>	<i>Coeff</i>	<i>z-value</i>	<i>Coeff</i>	<i>z-value</i>
Age	0.021	24.66	0.130	78.82	0.017	26.26
Male	-0.697	-26.98	-0.252	-6.14	-0.475	-23.1
Black	0.439	12.07	-0.100	-1.73	-0.551	-17.78
Hispanic	-0.478	-15.48	-0.949	-17.09	-1.442	-58.77
Othrace	0.214	3.55	-0.368	-3.41	-0.472	-9.35
Married	-0.301	-11.13	0.109	2.58	0.522	24.41
pl_100to200	-0.307	-9.90	0.148	2.96	0.322	11.9
pl_200to300	-0.670	-16.42	0.191	3.17	0.885	28.93
pl_300to400	-0.861	-15.33	0.354	4.73	1.369	37.39
pl_gt400	-0.505	-10.81	0.468	7.05	1.900	56.56
bad_hlth	0.324	3.68	2.232	12.03	-1.341	-15.1
lim_adl	1.841	13.93	0.918	6.45	0.969	7.31
urban_area	-0.009	-0.28	0.147	3.17	0.256	10.01
y97	-0.031	-0.27	-0.092	-0.51	-0.204	-2.43
y98	0.093	0.81	0.102	0.55	-0.260	-2.99
y99	1.272	6.27	0.640	1.96	-0.784	-4.8
y00	1.196	6.32	0.729	2.39	-0.637	-4.17
y01	1.083	6.23	0.657	2.35	-0.515	-3.73
y02	1.361	8.86	0.691	2.79	-0.162	-1.34
y03	1.770	10.98	0.821	3.16	-0.027	-0.21
reg_firm_ins_off	2.900	5.49	1.336	1.67	1.599	3.57
reg_hi_cont	0.000	-1.17	0.000	-1.74	-0.002	-13.53
reg_aa_mort	-0.003	-8.08	-0.001	-1.63	0.002	5.13
age_badhlth	0.007	3.95	-0.024	-8.3	0.021	12.29
Intercept	-0.065	-0.17	-8.149	-13.88	1.282	4.15
<i>N</i>	91,039					
Log likelihood	-71,432					
LR Test χ^2_{72}	31,165					
Pseudo R^2	0.231					

Note: standard errors have been bootstrapped and made robust to account for clustered regressors.

would expect health insurers to regionally adjust premiums and cost-sharing requirements according to expected mortality risks. We obtain a log pseudolikelihood value of -71,432 when we include the instruments in the insurance regression, and the correlation between these instruments and health expenditure was insignificant.

In terms of the sample selection equation, two instruments are employed. The first instrument is the regional concentration of doctors, which may be a supply-side factor that influences an individual's contact with the health care system. The second instrument is total regional health expenditure, which may be reflective of access to care within the region. The log pseudolikelihood value is -43,255, and again, the correlation between these instruments and health expenditure was insignificant.

The determinants of insurance coverage

Table 2 presents the results from the insurance equations. The dependent variable was categorized as Medicare only, other public insurance, private insurance, and no insurance, and the base outcome in Equation (1) was: “does not have insurance”. To avoid multicollinearity problems, we checked the condition index of the model. Overall, the diagnostic indicators suggest an acceptable goodness of the fit (pseudo- R^2 of 0.231). Age has the same effect on all types of insurance coverage, indicating that compared with being uninsured, the probability of having Medicare only, other public, or private insurance increases with age. Males are less likely to be publicly or privately insured, possibly due to being less risk averse than women.⁴⁸ As expected, race/ethnicity is an important predictor of insurance: blacks are more likely to obtain other public insurance than whites, while blacks and Hispanics are less likely to have private insurance than whites. We categorized the income variable by the Federal poverty level (\$7,955 for one individual in 1996), although education was excluded due to its multicollinearity with income. As expected, higher income individuals are less likely to have public insurance (excluding Medicare only and Medicare and private insurance) and more likely to have private insurance or Medicare only.

We find that poor health status is positively associated with public insurance and negatively associated with private insurance,⁴⁹ possibly indicating the presence of risk selection. However, individuals with at least one ADL are more likely to have any form of insurance. Living in an urban area is associated with a higher probability of having Medicare only or private insurance, potentially due to a greater density of insurers and providers in urban areas. Finally, the year controls, which are expected to capture time-specific changes in insurance conditions, are significant.

Effects on the probability of any health expenditure

The results of the probit regression for any health expenditure are listed in Table 3. Interestingly, age is not a significant predictor of any health-care use, possibly because individuals of all ages are likely to have some contact with the health care system. Consistent with prior evidence, women are more likely to incur any form of expenditure. Married individuals also have a lower likelihood of any expenditure, perhaps because spouses have to share income across at least two individuals. There is a negative relationship between income and the probability of expenditure, potentially because the primary route of income is through the choice of insurance. Individuals in poor health or with physical impairments (ADLs) consistently have a higher likelihood of expenditure. The effect of insurance is as expected, with individuals who have Medicare only or private health insurance having a higher likelihood of any health expenditure, while those with other forms of public insurance exhibiting lower probabilities of health-care consumption, possibly because of reduced access to care.

⁴⁸ Zinkhan and Karande (1991); Barsky *et al.* (1995).

⁴⁹ Hurd and McGarry (1997).

Table 3 Probability of any health expenditure^a (1996–2003)

<i>Variable</i>	<i>Health expenditures</i>	
	<i>Coeff.</i>	<i>z-value</i>
age	−0.003	−1.19
male	−0.614	−31.61
married	−0.195	−13.28
pl_100to200	−0.255	−15.00
pl_200to300	−0.512	−20.34
pl_300to400	−0.706	−23.87
pl_gt400	−0.661	−20.83
bad_hlth	0.778	22.01
lim_adl	0.812	8.84
mcr_only_hat	−0.290	−12.26
oth_pub_hat	0.166	6.54
any_prv_hat	0.345	26.17
urban_area	−0.162	−13.53
reg_conc_drs	0.000	1.69
reg_hlth_exp	4.530	2.85
y97	−0.061	−1.68
y98	0.027	0.76
y99	−0.109	−2.73
y00	−0.092	−2.24
y01	−0.056	−1.27
y02	0.038	0.99
y03	0.146	3.63
Intercept	0.892	3.00
<i>N</i>	91,039	
Wald (22)	30,300	
Pseudo		
Log-Likelihood	−43,255	

Note: standard errors have been bootstrapped and made robust to account for clustered regressors.

^aOn prescription drugs, inpatient stays, outpatient visits, emergency room visits, or office-based visits.

Effects on expenditure

Table 4 reports the expenditure estimates after correcting for endogeneity and sample selection. The sample selection coefficient (inverse Mills ratio) is significant for all sources of expenditure except inpatient care, suggesting that the selection correction influences the estimated coefficient. With the exception of inpatient care, individuals with Medicare only and individuals with private health insurance spend more on all types of health care than individuals with no insurance, while individuals with other public insurance spend less on all types of health care. For pharmaceutical care, there is a positive age effect, in that elderly patients spend more than non-elderly patients, males spend less than females, married individuals spend less than non-married persons, and individuals with incomes higher than the poverty level spend less than the poorest individuals. Medicare and private insurees also pay more for prescription medicines than the uninsured. For inpatient care, respondents with other public

Table 4 Multivariate regression of all components of health expenditure (1996–2003)

Variable	Pharmaceutical		Outpatient		Inpatient		Emergency		Office visit	
	Coeff	z-value	Coeff	z-value	Coeff	z-value	Coeff	z-value	Coeff	z-value
age	0.006	1.96	-0.044	-8.94	0.055	7.71	-0.040	-7.76	-0.003	-0.72
male	-0.072	-2.65	-0.368	-5.90	-0.009	-0.11	-0.154	-3.38	-0.131	-2.74
married	-0.272	-17.55	-0.164	-3.99	-0.065	-1.65	-0.204	-8.20	-0.031	-1.13
pl_100to200	-0.247	-8.71	-0.328	-9.47	-0.072	-1.30	-0.044	-1.32	-0.053	-1.74
pl_200to300	-0.438	-13.04	-0.538	-8.08	-0.381	-4.72	-0.123	-3.43	-0.077	-1.47
pl_300to400	-0.588	-12.37	-0.824	-9.76	-0.442	-3.92	-0.224	-4.48	-0.139	-1.91
pl_gt400	-0.582	-12.59	-0.941	-12.78	-0.672	-6.32	-0.299	-5.54	-0.158	-2.48
bad_hlth	0.633	22.02	0.509	8.60	1.341	16.31	0.476	9.56	0.315	5.51
lim_adl	0.674	15.09	-0.182	-1.78	1.616	11.17	0.620	9.18	-0.001	-0.01
mcr_only_hat	-0.347	-14.94	-0.303	-6.67	0.046	0.75	-0.064	-1.86	-0.093	-2.28
oth_pub_hat	0.210	7.22	0.545	11.02	-0.453	-6.55	0.327	6.67	0.148	3.33
any_prv_hat	0.076	4.1	0.345	11.07	0.268	5.41	0.013	0.50	0.092	3.06
urban_area	-0.048	-2.97	-0.348	-12.97	-0.073	-2.53	-0.142	-6.82	0.114	4.96
y97	0.180	4.73	0.025	0.31	0.019	0.20	0.055	0.83	-0.030	-0.44
y98	0.275	6.55	0.059	0.70	0.114	1.20	-0.005	-0.07	-0.039	-0.54
y99	0.330	7.66	-0.154	-1.93	0.124	1.37	-0.078	-1.10	0.082	1.31
y00	0.277	6.13	-0.199	-2.29	0.158	1.81	-0.068	-0.96	0.029	0.44
y01	0.489	10.61	-0.085	-0.98	0.098	0.97	0.017	0.25	-0.011	-0.15
y02	0.650	16.64	0.240	2.95	0.113	1.37	0.110	1.66	0.084	1.35
y03	0.830	20.97	0.378	4.26	0.092	0.96	0.175	2.53	0.221	3.43
health_mills	-1.396	-13.91	0.645	3.43	-0.403	-1.53	1.030	6.97	-1.624	-10.67
intercept	5.768	25.75	4.472	12.28	-2.816	-5.56	3.345	8.69	5.861	16.99
F-Test	1,046		163		155		73		311	
R ²	0.274		0.056		0.053		0.026		0.101	
P	0.000		0.000		0.000		0.000		0.000	

Note: standard errors have been bootstrapped.

insurance or private insurance spend more than uninsured individuals, while respondents with Medicare only spend less, although this coefficient was not significant.

There are a number of reasons why we might have observed these patterns. Although the poorest groups in the population tend to be less healthy than the rest of the population,⁵⁰ it appears that the poor may experience barriers in access to care, for instance, these individuals may not be able to afford to take time off from work for a few hours or days. Medicaid also pays lower prices for medical care than uninsured individuals because of strong negotiating power. The fact that individuals with Medicare only spend more on most types of care than the uninsured is not surprising, given that this population group is older, less healthy, and does not face the full price of care for most outpatient services. This group even spends more on prescription drugs even though Medicare did not offer an

⁵⁰ Macinko *et al.* (2003).

outpatient drug benefit during the study period, although this outcome was likely related to greater need for medications in this population. Yet, Medicare individuals incur lower inpatient expenditure than uninsured persons, perhaps because of Medicare's pricing power and because Medicare enrollees may see primary prevention as more important. Those with private insurance spend more on all types of medical care, even after accounting for health status and income. Thus, there appears to be a significant moral hazard effect associated with private insurance, and this volume effect generally outweighs the effect of price negotiations.

Both drug and inpatient expenditure increase with age, which is consistent with the preliminary evidence and the expected effect of ageing. However, age has the opposite effect on outpatient and emergency room expenditure, while it is not significant for office-based visits. Therefore, although ageing exerts an effect on expenditure, the effect is heterogeneously different across various sources of expenditure, suggesting that heterogeneity should be taken into account when examining expenditure data. Across all types of expenditure (except for inpatient), males and married individuals exhibit lower expenditure. After controlling for the endogeneity of insurance, the effect of income is systematically negative, indicating that higher income does not lead to higher health expenditure. The likely explanation is that higher-income individuals purchase more generous insurance, and the increase in consumption is channelled through the insurance effect rather than an income effect. Meanwhile, individuals living in an urban area spend less on all types of care, with the exception of office-based care. This conflicting result for office-based care might be related to a higher concentration of doctors in urban areas.

As expected, poor health status is always associated with higher expenditure, and although having an ADL limitation leads to lower expenditure for outpatient and office-based visits, neither of these results is significant. Finally, the time effects are consistent with the aggregate evidence, suggesting that expenditure peaks in 2002–2003 and indicating that time effects are related to the expansion of drug and outpatient expenditure.

Discussion

By using several years of micro data, the intent of this paper has been to analyse the effects of insurance and need-related determinants on pharmaceutical and other health expenditure and to explain the causes behind recent health expenditure growth in the U.S. In addressing the research questions, we have corrected for three issues: (i) the potential endogeneity of insurance, (ii) the joint determination of different types of health expenditure, and (iii) unobserved health care consumption for some individuals (sample selection). As hypothesized, the corrections suggest that there is an heterogeneous effect of insurance on health care expenditure. We find that the joint estimation of health expenditure as well as additional controls, such as endogeneity adjustments for insurance and sample selection corrections, influence the estimates for each type of expenditure. Furthermore, it appears that

socio-economic, demographic, and health characteristics have heterogeneous effects on different components of health care expenditure. Yet, it is important to bear in mind that we are unable to fully capture the unobserved heterogeneity in health expenditure related to specific fixed effects. Future research could address this methodological issue.

The results obtained in this study have a number of policy implications. First, with the exception of inpatient care, age exhibits a non-linear pattern consistent with other studies, indicating some accumulation of health expenditure at early and later stages of life.⁵¹ Accordingly, policymakers should be less concerned about the impact of an ageing population on health care expenditure and more concerned about other factors behind these costs. Policies aimed at reducing expenditure among the elderly and any proposals to increase premiums for the higher-income elderly may end up doing more harm than good and detract from the real drivers of expenditure increases.

Second, although drug expenditure is the primary source of overall health expenditure increases, it is not necessarily related to the expansion of public insurance coverage. Perhaps there are barriers to access among public insurance beneficiaries, or perhaps Medicaid programmes are relatively successful at containing expenditure increases. Nonetheless, there is evidence of *ex-post* moral hazard among the privately insured population, explaining the consumption of more expensive treatments rather than determining use. The latter is an important finding considering that these individuals are likely to be the healthier and wealthier Americans. This suggests that measures implemented by private insurance companies have not been completely effective at containing expenditure. The nature of the data does not permit us to determine the reasons for these insurance and expenditure-related findings. However, these results do imply that attempts to privatize Medicare and Medicaid and proposals that involve expanding health insurance coverage through tax credits may do little to contain health expenditure, and in fact, may only lead to higher costs. A less costly option may actually be to expand public insurance coverage, while ensuring that public programmes continue to impose tight reimbursement rules. Of course, there may be health and access implications from imposing these restrictions, but these considerations are outside the scope of the present study.

Third, after controlling for the endogeneity of insurance and sample selection, it appears that income has a negative impact on most expenditure. This may have occurred because health status was included in the analysis, and higher income tends to be positively correlated with health status. Another reason for this finding is that the primary effect of income appears to be through the purchase of insurance. However, the generally negative impact of income on expenditure does not imply that insurance coverage should not be targeted at the poorest groups in the population; on the contrary, it suggests that policymakers need to ensure that these individuals are covered as they are less likely to purchase private insurance and are also more likely to be in poorer health. The recent Federal bill aimed at

⁵¹ Zweifel *et al.* (1999, 2004).

increasing cost sharing for the Medicaid population⁵² is an example of a policy that will likely shift costs to the most vulnerable in the population and potentially reduce allocative efficiency.

Acknowledgements

The authors are grateful for the comments of two anonymous referees on an earlier draft. All outstanding errors are the authors' own.

References

- Agency for Healthcare Research Quality (AHRQ) (2004) *MEPS HC-060: 2001 Full Year Consolidated Data File*, Rockville, MD: Agency for Healthcare Research and Quality.
- Amemiya, T. (1974) 'Multivariate regression and simultaneous equation models when the dependent variables are truncated normal', *Econometrica* **42**: 999–1012.
- Angrist, J.D. and Krueger, A.B. (2001) 'Instrumental variables and the search for identification: From supply and demand to natural experiments', *Journal of Economic Perspectives* **15**(4): 69–85.
- Artz, M., Hadsall, R.S. and Schondelmeyer, S.W. (2002) 'Impact of generosity level of outpatient prescription drug coverage on prescription drug events and expenditure among older persons', *American Journal of Public Health* **92**: 1257–1263.
- Barsky, R.F., Kimball, M. and Shapiro, M. (1995) *An experimental approach to preference parameters and behavioral heterogeneity in the health and retirement study*, Health and Retirement Study Working Paper Series, Paper No. 94-019.
- Cartwright, W., Hu, T. and Huang, L. (1992) 'Impact of varying Medigap insurance coverage on the use of medical services of the elderly', *Applied Economics* **24**: 529–539.
- Centers for Medicare and Medicaid Services (CMS) (2006) 'CMS Programs and Information', from <http://www.cms.hhs.gov/> accessed 26 January 2006.
- Christensen, S., Long, S.H. and Rodgers, J. (1987) 'Acute health care costs for the aged Medicare population: Overview and policy options', *Milbank Quarterly* **65**: 397–425.
- Contayannis, P., Hurley, J., Grootendorst, P., Jeon, S.H. and Tamblyn, R. (2005) 'Estimating the price elasticity of expenditure for prescription drugs in the presence of nonlinear price schedules: An illustration from Quebec, Canada', *Health Economics* **14**: 909–923.
- Duan, N., Manning, W.G., Morris, C.N. and Newhouse, J.P. (1983) 'A comparison of alternative models for the demand for medical care', *Journal of Economic and Business Statistics* **1**(2): 115–126.
- Federman, A., Adams, A.S., Ross-Degnan, D., Soumerai, S.B. and Ayanian, J.Z. (2001) 'Supplemental insurance and use of effective cardiovascular drugs among elderly Medicare beneficiaries with coronary heart disease', *Journal of the American Medical Association* **286**: 1732–1739.
- Feldman, R., Finch, M., Dowd, B. and Cassou, S. (1989) 'The demand for employment-based health insurance plans', *The Journal of Human Resources* **24**: 115–142.
- Frank, R.G. (2001) 'Prescription drug prices: Why do some pay more than others do?' *Health Affairs* **20**(2): 115–128.
- Fronstin, P., Goldberg, L. and Robins, P. (1997) 'An analysis of the decline in private health insurance coverage between 1988 and 1992', *Social Science Quarterly* **78**: 44–65.
- Gabel, J., Dhont, K., Whitmore, H. and Pickreign, J. (2002) 'Individual insurance: How much financial protection does it provide?', *Health Affairs Web Exclusive* **21**(3): W172–W181.
- Heckman, J.J. (1979) 'Sample selection bias as a specification error', *Econometrica* **47**: 153–161.
- Hurd, M.D. and McGarry, K. (1997) 'Medical insurance and the use of health care services by the elderly', *Journal of Health Economics* **16**: 129–154.

⁵² KFF (2006a).

- Huskamp, H., Deverka, P.A., Epstein, A.M., Epstein, R.S., McGuigan, K.A. and Frank, R.G. (2003) 'The effect of incentive-based formularies on prescription-drug utilization and spending', *New England Journal of Medicine* **349**: 2224–2232.
- Jones, A. (2001) 'Health econometrics', in A.J. Culyer and J.P. Newhouse (eds) *Handbook of Health Economics, Volume 1*, Amsterdam: Elsevier, pp. 267–344.
- Joyce, G., Escarce, J.J., Solomon, M.D. and Goldman, D.P. (2002) 'Employer drug benefit plans and spending on prescription drugs', *Journal of the American Medical Association* **288**: 1733–1739.
- Kaiser Family Foundation (KFF) (2006a) 'Deficit Reduction Act of 2005: Implications for Medicaid', from <http://www.kff.org/medicaid/upload/7465.pdf>, accessed 25 June 2006.
- Kaiser Family Foundation (KFF) (2006b) 'Trends and Indicators in the Changing Health Care Marketplace', from <http://www.kff.org/insurance/7031/print-secl.cfm>, accessed 14 June 2006.
- Kaiser Family Foundation and Health Research and Educational Trust (KFF/HRET) (2005) 'Employer Health Benefits: 2005 Annual Survey', from <http://www.kff.org/insurance/7315/upload/7315.pdf>, accessed 14 June 2006.
- Kanavos, P. and Gemmill, M. (2004) 'Senior citizens and the burden of prescription drug outlays: What lessons for the Medicare prescription drug benefit?' *Applied Health Economics and Health Policy* **3**: 217–227.
- Leung, S.F. and Yu, S. (1996) 'On the choice between sample selection and two-part models', *Journal of Econometrics* **72**: 197–229.
- Liebowitz, A., Manning, W. and Newhouse, J. (1985) 'The demand for prescription drugs as a function of cost-sharing', *Social Science and Medicine* **21**: 1063–1069.
- Lillard, L., Rogowski, J. and Kington, R. (1999) 'Insurance coverage for prescription drugs: Effects on use and expenditures in the Medicare population', *Medical Care* **37**: 926–936.
- Long, S. (1994) 'Prescription drugs and the elderly: Issues and options', *Health Affairs* **13**(2): 157–174.
- Macinko, J., Shi, L., Starfield, B. and Wulu, J.T. (2003) 'Income inequality and health: A critical review of the literature', *Medical Care Research and Review* **60**: 407–452.
- Maddala, G.S. (1985) 'A survey of the literature on selectivity bias as it pertains to health care markets', in R.M. Scheffler and L.F. Rossiter (eds) *Advances in Health Economics and Health Services Research, Volume 6*, Greenwich: JAI Press, pp. 3–17.
- Manning, W., Morris, C.N., Newhouse, J.P., Orr, L.L., Duan, N., Keeler, E.B., Liebowitz, A., Marquis, K.H., Marquis, M.S. and Phelps, C.E. (1981) 'A two-part model of the demand for medical care: Preliminary results from the health insurance study', in J. Van der Gaag and M. Perlman (eds) *Health, Economics, and Health Economics*, Amsterdam: North-Holland Publishing Company, pp. 103–123.
- Mills, J.A. and Zandvakili, S. (1997) 'Statistical inference via bootstrapping for measures of inequality', *Journal of Applied Econometrics* **12**: 133–150.
- Moulton, B.R. (1986) 'Random group effects and the precision of regression estimates', *Journal of Econometrics* **32**: 385–397.
- Newhouse, J. and The Insurance Experiment Group (1993) *Free for All? Lessons from the RAND Health Insurance Experiment*, Cambridge, MA: Harvard University Press.
- Newhouse, J., Manning, W.G., Morris, C.N., Orr, L.L., Duan, N., Keeler, E.B., Liebowitz, A., Marquis, K.H., Marquis, M.S., Phelps, C.E. and Brook, R.H. (1981) 'Some interim results from a controlled trial of cost sharing in health insurance', *New England Journal of Medicine* **305**: 1501–1507.
- Organisation for Economic Co-operation and Development (OECD) (2005) 'OECD Health Data 2005', from http://www.oecd.org/document/56/0,2340,en_2649_34631_12968734_1_1_1_1,00.html, accessed 12 Jan 2006.
- Pauly, M. (1974) 'Overinsurance and public provision of insurance: The roles of moral hazard and adverse selection', *The Quarterly Journal of Economics* **88**: 44–62.
- Skinner, J. and Wennberg, J.E. (2000) 'Regional inequality in Medicare spending: The key to Medicare reform?', *Frontiers in Health Policy Research* **3**(1): 67–90.
- Smith, D. (1993) 'The effects of copayments and generic substitution on the use and costs of prescription drugs', *Inquiry* **30**: 189–198.

- Smith, R. (1987) 'Testing for exogeneity in limited dependent variable models using a simplified likelihood ratio statistic', *Journal of Applied Econometrics* **2**: 237–245.
- Street, A., Jones, A. and Furuta, A. (1999) 'Cost-sharing and pharmaceutical utilisation and expenditure in Russia', *Journal of Health Economics* **18**: 459–472.
- The White House (2006) 'State of the Union: Affordable and Accessible Health Care', from <http://www.whitehouse.gov/news/releases/2006/01/20060131-7.html>, accessed 30 January 2006.
- U.S. Census Bureau (2005) 'Income, Poverty, and Health Insurance Coverage in the United States', 2004 from <http://www.census.gov/prod/2005pubs/p60-229.pdf>, accessed 14 June 2006.
- U.S. Department of Commerce (2006) 'Statistical Abstract of the United States', from <http://www.census.gov/compendia/statab/>, accessed 30 January 2006.
- Vella, F. (1998) 'Estimating models with sample selection bias: A survey', *Journal of Human Resources* **33**: 127–172.
- White, H. (1980) 'A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity', *Econometrica* **48**: 817–830.
- Wooldridge, J.M. (2002) *Econometric Analysis of Cross Section and Panel Data*, Cambridge, MA: The MIT Press.
- Yang, Z., Gilleskie, D. and Norton, E. (2004) *Prescription Drugs, Medical Care, and Health Outcomes: A Model of Elderly Health Dynamics*, Cambridge, MA: National Bureau of Economic Research.
- Zinkhan, G.M. and Karande, K.W. (1991) 'Cultural and gender differences in risk-taking behavior among American and Spanish decision makers', *The Journal of Social Psychology* **131**: 741–742.
- Zweifel, P., Felder, S. and Meiers, M. (1999) 'Ageing of the population and health care expenditure: A red herring?' *Health Economics* **8**: 485–496.
- Zweifel, P., Felder, S. and Werblow, A. (2004) 'Population ageing and health care expenditure: New evidence on the "red herring"', *The Geneva Papers on Risk and Insurance — Issues and Practice* **29**: 652–666.

Appendix

Table A1 Expected impact of explanatory variables on prescription drug expenditure

<i>Variable</i>	<i>Expenditure</i>	<i>Studies</i>
Insurance coverage	+	Contayannis <i>et al.</i> (2005), Liebowitz <i>et al.</i> (1985), Smith (1993), Joyce <i>et al.</i> (2002), Huskamp <i>et al.</i> (2003)
Coverage (vs.none)	+	Artz <i>et al.</i> (2002)
Medicaid	–	Yang <i>et al.</i> (2004)
Coverage for prescription drugs (vs. no supplemental coverage)	+	Long (1994), Federman <i>et al.</i> (2001), Yang <i>et al.</i> (2004)
Female	+	Liebowitz <i>et al.</i> (1985)
Age	–/+	Lillard <i>et al.</i> (1999), Contayannis <i>et al.</i> (2005)
Household size	+	Smith (1993), Street <i>et al.</i> (1999)
Urban area	+	Lillard <i>et al.</i> (1999)
Poor health	+	Street <i>et al.</i> (1999)
Income	+	Hurd and McGarry (1997)
Health	+	Hurd and McGarry (1997)

Table A2 Main variables: definitions and summary statistics

<i>Variable</i>	<i>Definition</i>	<i>Mean</i>	<i>Standard error</i>
<i>Endogenous variables</i>			
ln_rx_spend	Natural log of prescription drug expenditure	5.438	0.007
ln_houtp_spend	Natural log of hospital outpatient visit (physician and non-physician) expenditure	0.963	0.008
ln_hinp_spend	Natural log of hospital inpatient stay (including zero night stays) expenditure	0.716	0.008
ln_erspend	Natural log of emergency room visit expenditure	0.632	0.006
ln_office_spend	Natural log of office-based visit (physician and non-physician) expenditure	3.780	0.010
<i>Explanatory variables</i>			
<i>Demographics</i>			
Age	Age of the individual	44.49	0.058
Male	Individual is male	0.463	0.002
Urban_area	Individual lives in an urban area	0.276	0.001
<i>Socio-economic</i>			
Married	Individual is married	0.564	0.002
pl_lt100	Individual has an income less than 100 per cent of the FPL	0.259	0.001
pl_100to200	Individual has an income between 100 per cent and 200 per cent of the FPL	0.197	0.001
pl_200to300	Individual has an income between 200 per cent and 300 per cent of the FPL	0.150	0.001
pl_300to400	Individual has an income between 300 per cent and 400 per cent of the FPL	0.118	0.001
pl_gt400	Individual has an income greater than 400 per cent of the FPL	0.276	0.001
<i>Health status</i>			
bad_hlth	Health of individual is rated as poor or fair	0.142	0.001
lim_adl	Individual faces at least one limitations to an activity of daily living	0.022	0.000
<i>Insurance status</i>			
mcr_only	Individual has Medicare and no other insurance coverage	0.141	0.001
oth_pub	Individual has public insurance (Medicare with another form of public coverage, Medicaid, TRICARE, VA coverage, or other public insurance)	0.053	0.001
any_prv	Individual has private insurance	0.636	0.002

About the Authors

Marin Gemmill is Merck Trust Foundation Fellow at the London School of Economics, and she is working on her Ph.D in Health Economics and Policy at LSE. She holds an M.A. in Economics from the University of Delaware and an M.Sc. in International Health Policy from LSE.

Joan Costa-Font is Associate Professor of Economics at the University of Barcelona and Research Fellow in Health and Pharmaceutical Economics at the London School of Economics.

Panos Kanavos is Lecturer in international health policy and Research Fellow in Health and Pharmaceutical Economics at the London School of Economics. He is also head of the Pharmaceutical Economics and Policy Group at LSE Health.