Determinants of healthcare spending: a state level analysis

James A. Thornton a & Jennifer L. Rice a

a Department of Economics, Eastern Michigan University, Ypsilanti, Michigan 48197, USA

Published online: 11 Apr 2011.
Determinants of healthcare spending: a state level analysis

James A. Thornton and Jennifer L. Rice*

Department of Economics, Eastern Michigan University, Ypsilanti, Michigan 48197, USA

Within the high and rising level of healthcare spending for the US as a whole is substantial variation in spending across states. Yet relatively little attention has been given to the empirical analysis of interstate differences in aggregate healthcare expenditures, and therefore little empirical evidence exists at the state level to guide policymakers. Using data for all 50 states for the year 1998, we estimate an empirical model that includes structural and reduced-form healthcare spending equations and a health production function to assess the significance, size and relative importance of factors that prior research indicates, may play an important role in explaining interstate variation in medical care expenditures, and the main pathways through which they operate. Our results indicate higher levels of healthcare spending for state populations with higher income, less education, fewer uninsured residents, less healthy lifestyles, larger proportion of elderly residents, greater availability of medical care providers and less urbanization. Our findings suggest that the most effective cost containment measures may be those that increase education and promote healthy lifestyles. Not only do these actions lead to reductions in healthcare spending, they also improve the health status of the population, and may help to achieve other important social policy goals.

I. Introduction

Over the past several decades medical care expenditures in the US have risen rapidly, increasing from $215 billion in 1980 to $1.02 trillion in 1998, an average annual growth rate of 9%. By 2003, annual spending on medical care reached $1.7 trillion, consuming 15% of the gross domestic product (Smith et al., 2005). Within the high and rising level of healthcare expenditures for the US as a whole is substantial variation in spending across states. For example, in 1998 medical care spending per capita ranged between a low of $2731 in Utah and a high of $4810 in Massachusetts (Martin et al., 2002). Many argue that the most important factor explaining high and rising healthcare spending is advancements in medical technology (Newhouse, 1992; Kleinke, 2001; Glied, 2003). However, because medical knowledge and new interventions tend to flow rapidly over geographic regions, technology cannot account for the dramatic differences in spending across states (Fuchs and Kramer, 1980; Fuchs, 2004). An understanding of nontechnology factors that explain interstate variation in healthcare expenditures would help to illuminate determinants of healthcare costs and assist policymakers in designing effective cost containment strategies.

A large number of studies have used multivariate regression models to explain differences in aggregate healthcare spending across countries...
(Gerdtham and Jonsson, 2000). In general, these studies find that aggregate income explains a
preponderance of the variation in international healthcare expenditures with an income elasticity of one or
greater. Other factors that appear to be important in
cross-country comparisons include type of insurance
arrangement, supply of physicians, use of primary
care gatekeepers and public provision of healthcare.
Most studies of aggregate healthcare spending in the
US have attempted to estimate the total effect of
selected factors, such as population age structure,
medical malpractice insurance, medical errors and
administrative costs on medical care costs. More
recently, empirical research has started to focus on
the impact of obesity and other determinants of
health status on healthcare expenditures. These
studies suggest that health related lifestyle factors
are important determinants of healthcare spending,
with estimates of the annual cost of overweight and
obesity as high as $93 billion (Finkelstein et al., 2003).
While these total cost estimates are informative,
policymakers designing effective cost containment
strategies could benefit by having knowledge about
factors that significantly affect medical care spending
across geographic regions within the US, as well as
the relative size of their marginal effects. However,
relatively little attention has been given to the
empirical analysis of interstate differences in aggre-
gate healthcare spending, and therefore surprisingly
little empirical evidence exists at the state level to
guide policymakers.

The objective of our article is to analyse factors
that influence healthcare expenditures across states.
Our goal is to identify the most important cross-
section determinants of spending on medical care
services, and obtain estimates of the relative size of
their effects. We are particularly interested in the
effect on spending of health related determinants such
as lifestyle and socio-economic factors, and the
potential pathways through which they operate. In
spite of mounting evidence that health related factors
are an important force explaining medical care
spending, the significance and size of their effects
have remained largely unexplored in multivariate
analyses of aggregate healthcare spending.

II. Analytical Framework

In this section we summarize the simplified analytical
framework, presented formally in the Appendix, we
use to guide specification of our empirical models and
interpret results. This framework describes the inter-
relationship between aggregate healthcare spending
and population health status, and illuminates potential
pathways through which health and health related
determinants affect spending.

Aggregate healthcare spending (S) is defined by
the identity $S = PQ$, where $P$ is the unit price of
healthcare and $Q$ is the quantity utilized. We assume
healthcare spending is the outcome of a demand and
supply process that determines the equilibrium price
and quantity of healthcare. However, in our frame-
work, equilibrium price and quantity are conditional
on population health status, which in turn depends
upon utilization of healthcare services. This implies
reduced-form and structural spending equations, and
a health production function, given by

\[
\ln S = \pi_0 + \pi_X \ln X + \pi_Y \ln Y + \pi_Z \ln Z \quad (1)
\]

\[
\ln S = \gamma_0 + \gamma_X \ln X + \gamma_Y \ln Z + \gamma_H \ln H \quad (2)
\]

\[
\ln H = \beta_0 + \beta_Y \ln Y + \beta_Z \ln Z + \beta_S \ln S \quad (3)
\]

where $\ln$ denotes natural logarithm, $H$ is population
health status, $X$ are exogenous variables that have a
direct effect on spending (e.g. health insurance,
competition among providers), $Y$ are exogenous
variables that have an indirect effect on spending by
affecting health status (e.g. age distribution, obesity)
and $Z$ are exogenous variables that have both direct
and indirect effects on spending (e.g. income, educa-
tion). Following the literature on the empirical
aggregate health production function, healthcare
spending (S) is used as a proxy for the utilization of
healthcare services ($Q$) in the health production
function, because state healthcare price indices do
not exist to obtain an estimate of the quantity of real
healthcare services by state.

Exogenous variables in $X$ and $Z$ include both
demand-side and supply-side variables that have a
direct effect ($X$), as well as variables with both a direct
and indirect effect ($Z$) on spending. We expect
demand factors to have a positive relationship with
price, utilization and spending. Assuming an inelastic
demand for healthcare, which is consistent with prior
empirical studies, we expect supply factors to have a
positive relationship with utilization and a negative
relationship with price and spending. We expect
exogenous variables in $Y$ that improve (reduce) health
to decrease (increase) demand for medical care, and
therefore decrease (increase) price, utilization and
spending.

The reduced-form spending coefficients, $\pi_i$, measure
the total (marginal) effect of a variable on
healthcare expenditures, which includes the direct
and/or indirect effect of the variable, as well as the
feedback effect on spending that results from the
spending-health interaction produced by the variable.
The structural spending coefficients, $\gamma_i$, measure the
Determinants of healthcare spending

III. Variable Specification and Data

This section discusses the variables included in the healthcare spending equations and provides a description of the data, which consists of 50 states for the year 1998, the most recent year for which the data used in the analysis is available. Descriptive statistics are presented in Tables A1 and A2 which provides a detailed list of data sources.

Healthcare spending

The variable to be explained is aggregate state healthcare expenditures per capita (S) defined as annual real healthcare spending by state residents. The Centers for Medicare and Medicaid measures nominal personal spending for healthcare services and products by the state in which providers are located; that is, it measures the value of healthcare services produced by healthcare providers in a state, not the amount consumed by persons residing in a state, and does not eliminate variation in spending, resulting from differences in price levels among states. Border-crossing by state residents for healthcare services and price level differences distort measures of real healthcare spending, and therefore may lead to incorrect conclusions. Therefore, we adjust healthcare expenditures for border-crossing using the net flow index developed by Martin et al. (2002), and state price level differences using a state cost of living index developed by Leonard and Friar (Leonard et al., 1999).

Health status (H)

A potentially important factor affecting healthcare spending across states is differences in population health status. To measure health status in the structural spending equation and the health production function, we use the crude death rate (CDR) per 100,000 population and interpret this measure as an inverse index of health: states with lower mortality rates have healthier populations, demand less medical care, and therefore are expected to have lower levels of healthcare spending.

Variables with direct effect on healthcare spending (X)

Prior research suggests that health insurance coverage, availability of medical care providers, and various factors that influence the cost of providing healthcare can have direct and indirect effects on healthcare spending. For example, suppose that state per capita income (Z) increases. This is expected to have a direct positive effect on healthcare spending (γZ > 0) by relaxing the state level budget constraint, augmenting the population’s ability to pay for healthcare, and thereby increasing the demand for healthcare. However, if this higher level of income leads to an improvement in the state’s health status independent of healthcare (βZ > 0), as much research suggests, then this indirect effect will result in an offsetting reduction in healthcare spending. Assuming the aggregate marginal effect of healthcare utilization on population health status is positive (βZ > 0) feedback effects ensue, which reduce the magnitude of both the direct and indirect income effects. Theoretically, the total effect of a change in income on healthcare spending (πZ) can be positive, negative or zero depending on the relative magnitudes of the direct and indirect effects that work in opposite directions.

To analyse factors that influence healthcare spending across states, we employ the following approach. We begin by estimating the reduced-form healthcare spending Equation 1 separately, which yields measures of the total effects of factors that explain variation in expenditures across states. To obtain a better understanding of the pathways through which these factors operate, we then estimate the structural healthcare spending Equation 2 and the health production function (3) jointly. This provides information on direct and indirect (health) effects of these factors on spending.
healthcare services can have a direct effect on healthcare spending by affecting the demand and/or supply of healthcare services.

Health insurance can affect healthcare demand and spending by providing greater access to medical care services and reducing out-of-pocket price. We include two insurance variables in the reduced-form and structural spending equations: the percent of a state’s population with health insurance coverage \((INS)\) and the percent who are recipients of Medicaid \((MED)\). We expect states with more extensive health insurance coverage to have higher healthcare spending.

Availability of medical care providers can have a direct effect on state healthcare spending by influencing both supply and demand of medical care services. In a competitive market environment with inelastic demand for medical care, an increase in supply of physicians and hospital beds leads to lower healthcare expenditures. However, contrary to the typical market mechanism greater availability of medical care providers may lead to increased demand and ultimately higher spending. This demand effect may operate through a number of potential channels including supplier-induced demand, where physicians and hospitals recommend unnecessary services to maintain income or gain financially in the face of greater competition, and time-price induced demand resulting from reduced travel and waiting times in locations with greater availability of providers (Eastaugh, 1992; Frech, 1996). To measure availability of providers, we include measures of the number of physicians \((DOC)\) and hospital beds \((BED)\) per 100,000 population.

We include two supply-side factors, medical malpractice insurance premiums and Health Maintenance Organization \((HMO)\) penetration that may have a direct effect on healthcare spending by influencing the cost of providing services. Medical malpractice insurance premiums may influence healthcare spending by affecting medical practice costs and signalling exposure to malpractice risk thereby inducing physicians to practice defensive medicine. However, if physicians behave as if they maximize profit and are insured against malpractice claims, then changes in malpractice cost may primarily lead to changes in physician net income and the rationale for defensive medicine is less clear (Thornton, 1999). The most frequently cited estimate of insurance premiums and defensive medicine is for year 1984 and indicates that together they account for about 15% of spending on physician services (Reynolds et al., 1987). To measure medical malpractice insurance premiums \((MAL)\), we use data on average professional liability premiums for self-employed physicians in thousands of dollars. The American Medical Association reports data on average premiums for the 9 census divisions in the US, as well as 10 individual states within these divisions. For our measure of malpractice premiums, we use the data for the 10 individual states; as a proxy for the remaining 40 states we use census division averages for these states. We adjust nominal premiums for price level differences using the cost of living index of Leonard and Friar (Leonard et al., 1999).

Health Maintenance Organizations may reduce healthcare costs and spending by adopting more effective cost containment measures and increasing competition in those geographic markets in which they locate. The variable \(HMO\) is measured by the percent of the state’s population not covered by an \(HMO\), and therefore lower values of \(HMO\) indicate greater \(HMO\) penetration.\(^3\)

Variables with indirect effect on healthcare spending \((Y)\)

A large body of literature in economics, epidemiology and medical research has analysed the effect of lifestyle, socio-economic, environmental and demographic factors on population health.\(^4\) In general, these studies find evidence that factors such as obesity, smoking, drinking, exercise, aging, income and education are significant determinants of health status. More recently, greater recognition has been given to the indirect effect of lifestyle on healthcare spending through its effect on health status. Alcohol consumption, cigarette consumption and obesity are known causes of chronic illness and mortality, such as heart disease, cancer, liver disease and diabetes. These chronic diseases, in turn, have been implicated as important factors driving medical care spending. For example, using individual level data Sturm (2002) finds that obesity, smoking and problem drinking have significant effects on health status, and increases average annual spending on inpatient and ambulatory care by $395, $230 and $150. A few studies have attempted to estimate the total effect of

\(^3\) The variable \(HMO\) does not measure managed care penetration, which includes preferred provider organizations as well as \(HMOs\).

\(^4\) For example, refer (Auster et al., 1969; Silver, 1972; Fuchs, 1974; Hadley, 1982; Rosen and Taubman, 1982; Blair et al., 1995; Lee et al., 1995; Elo and Preston, 1996; Backlund et al., 1999; Calle et al., 1999; Deaton, 2002; Meyers et al., 2002; Thornton, 2002; Smith et al., 2005).
Determinants of healthcare spending

obesity and cigarette smoking on state and national healthcare spending by aggregating individual data to the state level and developing a national spending model.

Miller et al. (1998) finds that for the US as a whole cigarette smoking accounts for 11.8% of total healthcare expenditures, and that the proportion of spending attributed to smoking varies considerably among individual states, ranging from 6.6 to 14.1%. Finkelstein et al. (2004) find that 6% of total healthcare spending for the US is attributed to obesity, while variation at the state level ranges from 4 to 7%. Sedentary lifestyle, which is associated with overweight and obesity, has been identified as an independent factor affecting health, and therefore linked to medical care spending (Lee et al., 1995; Meyers et al., 2002). Research has long shown that deterioration in health associated with aging is related to higher medical costs, with per capita healthcare spending for the cohort age 65 and older more than three times the level for the cohort age 34–44 (Reinhardt, 2003).

To analyse the indirect effect of lifestyle related factors and aging on healthcare spending across states, we include measures of obesity (BMI), alcohol consumption (ALC), cigarette consumption (CIG), exercise (EXC), and the elderly population (A65) in the reduced-form spending equation and the health production function. Obesity is measured as the percent of the population age 18 and older with a body mass index of 30 or greater. Alcohol consumption is measured as the number of gallons of beer, wine and spirits sold per capita. Cigarette consumption is measured as the number of packs sold per capita and is estimated using data on annual gross tax revenues and federal and state cigarette tax per pack. Exercise is measured as the percent of the population age 18 and older who exercise at least five times per week, 30 minutes or more per session. Elderly population is measured as the percent of the population age 65 or older.

Variables with direct and indirect effect on healthcare spending (Z)

Prior research suggests that socio-economic variables can have both direct and indirect effects (via health) on healthcare spending. Three such variables are included in the structural and reduced-form healthcare spending equations, and the health production function: income (INC), distribution of income (DIST) and education (EDU). Income is measured as personal income per capita adjusted for state price level differences. Education is measured by the percent of the population age 25 and older who have at least a high school education.

The distribution of income is measured by the Gini coefficient, which can take a value between 0 and 1. A value of 0 indicates equality of income for each household in a state, while a value of 1 indicates a single household receives all state income. Values closer to one, therefore, indicate greater disparity in a state’s income distribution.

Prior studies find that the effect of income on healthcare spending depends upon the unit of observation, with small or negative income elasticities for individuals, and elasticities of one or greater at the national level. Refer to Getzen (2000) for summary of these studies. Getzen (2000) argues that the positive relationship between income elasticity and level of aggregation can be largely explained by the difference between individual budget constraints, which health insurance and government financing is designed to eliminate, and group level budget constraints that determine the amount, which can be spent on healthcare by all members collectively. State level income constraints are expected to be partially binding with income elasticities, falling between those for individuals and the nation as a whole.

Prior income elasticity estimates for physician and hospital services that use aggregate state level data range between about 0.1 and 0.9 (Feldstein, 1971; Fuchs and Kramer, 1980). To our knowledge, Freeman (2003) is the only study to estimate the income elasticity of aggregate state healthcare spending. Using panel data for the period 1966 to 1998, he reports income elasticity estimates of 0.82–0.84, but does not adjust healthcare spending for border-crossing or state price level differences, and includes income as the only explanatory variable.

Prior research has largely ignored the potential indirect effect of income on healthcare spending through population health status. A large body of research finds a positive relationship between income and health, which tends to be stronger at lower incomes but persists throughout the income distribution. While part of this relationship is explained by the effect of health on income, and associated lifestyle factors and medical care spending on health, recent research supports a potentially large causal effect that runs from income to health through mechanisms such as psychosocial stress, social participation and sense of control (Smith, 1999; Deaton, 2002; Marmot, 2002, 2004). Feldstein (1971) appears to be the only prior study to acknowledge the potential importance of this pathway, and finds suggestive evidence for an indirect health effect on state hospital expenditures.

A number of studies also find a positive association between income inequality and mortality across states and metropolitan areas in the US, and provide tenable causal interpretations involving material,
community and psychosocial processes (Kaplan et al., 1996; Kennedy et al., 1996; Ross et al., 2000; Mueller, 2002). This suggests that the distribution of income may affect healthcare spending through its effect on population health. Moreover, if preferences for healthcare vary across the distribution of income, then income inequality may also have a direct effect on healthcare spending. However, to our knowledge no prior studies have explored the relationship between income distribution and aggregate healthcare spending. However, to our knowledge no prior studies have explored the relationship between income distribution and aggregate healthcare spending.

Economic and epidemiological studies find that education has a large and significant effect on health at the individual, state and national level. Grossman (2000) concludes that years of education is the most important correlate of health, and this conclusion is robust to different measures of health as well as the level of aggregation of the analysis. Many economists argue that education has a direct causal effect by increasing the efficiency with which individuals and populations can produce health, which operates independent of income, lifestyle and unobservable characteristics that may be correlated with health and years of schooling. Rosen and Taubman (1982) and Berger and Leigh (1989), and others provide empirical evidence in support of this causal mechanism. While more years of formal education should lower healthcare spending indirectly through better health, a number of studies using individual level data suggest schooling may also have a countervailing direct positive effect on medical care expenditures by influencing attitudes about when to seek medical care. Escarce and Puffer (1997) and Kapur et al. (2004) report that Medicaid recipients with more formal education tend to spend more on healthcare. Additional support for a direct spending effect for nonMedicare persons is offered by Mutchler and Burr (1991), Hong and Kim (2000) and Finkelstein et al. (2004).

Additional variables with potential direct and indirect effects on healthcare spending, which we include in the spending equations and health production function, are measures of urbanization, race and physician specialty. Urbanization (URB), measured as the percent of the population that resides in a standard metropolitan statistical area, is a proxy for a collection of factors such as pollution, congestion and access to health care that may have a direct or indirect effect on healthcare spending. The potential effect of race on healthcare spending is measured by the percent of the population that is black (BLK), and is included to proxy for possible cultural differences in health status and preferences for seeking medical care. It has been argued that the specialty distribution of physicians may affect health status, and healthcare costs and spending, because nonprimary care physicians provide a different quality of care, are more inclined to use high cost medical technology, and receive higher fees for services rendered. To account for this possibility, we include the variable SPC measured as the percent of nonfederal physicians in nonprimary care specialties.

IV. Empirical Model

The empirical model consists of reduced-form and structural healthcare spending equations and a health production function corresponding to Equations 1 through 3, given by

\[
\ln S_t = \pi_0 + \pi_1 \ln INC_t + \pi_2 \ln DIST_t + \pi_3 \ln EDU_t + \pi_4 \ln INS_t + \pi_5 \ln DOC_t + \pi_6 \ln BED_t + \pi_7 \ln URB_t + \pi_9 \ln A65t + \pi_9 \ln CIG_t + \pi_{10} \ln ALC_t + \pi_{11} \ln BMI_t + \pi_{12} \ln CIG_t + \pi_{13} \ln SPC_t + \pi_{14} \ln HMO_t + \pi_{15} \ln MAL_t + \pi_{16} \ln MEDt + \pi_{17} \ln BLKt + \varepsilon_t \]

(4)

\[
\ln S_t = \gamma_0 + \gamma_1 \ln CDR_t + \gamma_2 \ln INC_t + \gamma_3 \ln DIST_t + \gamma_4 \ln EDU_t + \gamma_8 \ln INS_t + \gamma_9 \ln DOC_t + \gamma_10 \ln BED_t + \gamma_8 \ln URB_t + \gamma_9 \ln SPC_t + \gamma_{10} \ln HMO_t + \gamma_{11} \ln MAL_t + \gamma_{12} \ln MED_t + \gamma_{13} \ln BLKt + \mu_t \]

(5)

\[
\ln CDR_t = \beta_0 + \beta_1 \ln S_t + \beta_2 \ln INC_t + \beta_3 \ln DIST_t + \beta_4 \ln EDU_t + \beta_8 \ln SPC_t + \beta_9 \ln BLK_t + \beta_{10} \ln URB_t + \beta_8 \ln A65t + \beta_9 \ln CIG_t + \beta_{10} \ln ALC_t + \beta_{11} \ln BMI_t + \beta_{12} \ln CIG_t + \nu_t \]

(6)

where all variables are defined as above.

To account for heteroscedasticity of grouped state level data, we weight observations by the square root of state populations. We begin by estimating the reduced-form spending Equation 4 separately using ordinary least squares (OLS) estimation. To obtain unbiased and efficient estimates of structural parameters we then estimate the spending Equation 5 and health production function 6, jointly using a three-stage least squares (3SLS) estimation procedure, which accounts for endogeneity of health status (CDR) and spending (S), as well as possible correlation of disturbances across the two equations.

5 For example, refer (Auster et al., 1969; Silver, 1972; Fuchs, 1974; Grossman, 1975, 2000; Newhouse and Friedlander, 1980; Kenkel, 1991; Elo and Preston, 1996; Backlund et al., 1999; Thornton, 2002).
Determinants of healthcare spending

To identify the two structural equations, instruments must be found for health status in Equation 5 and healthcare spending in Equation 6. The analytical framework presented above indicates that a set of potential instruments for health status is provided by variables contained in the vector Y in the reduced-form spending equation, which have an indirect effect on healthcare spending only through their effect on population health status. We maintain that $A_{65}, CIG, ALC, BMI$ and $EXC$ satisfy this criterion, and therefore exclude these variables from the structural spending equation to obtain identification. A set of potential instruments for healthcare spending in the health production function include the variables $INS, MED, DOC, BED, HMO$ and $MAL$ that we maintain belong in the vector X. At last, we assess the validity of the specification of our equations by performing tests of instrument relevance, overidentifying restrictions, and variable specification and analyse sensitivity of estimates to alternative estimation procedures.

V. Empirical Results

Reduced-form healthcare spending equation

Table 1 presents the results of OLS estimation for the reduced-form spending equation. The $R^2$ statistic of 0.895 indicates that the model fits the data well, and suggests that the 17 factors included in Equation 4 explain a substantial amount of variation in healthcare spending across states. Recall that the reduced-form parameters measure the total (marginal) effect of a change in a factor on healthcare spending as an elasticity. A majority of the parameter estimates have expected signs, plausible magnitudes and reasonable significance levels. The exceptions are $DIST, SPC, BLK, HMO, MED$ and $MAL$. However, for most of these variables the coefficient estimates are very small, and all of these estimates are highly insignificant with t-statistics ranging from 0.12–1.21. An F-test of the joint significance of these six variables yields an F-statistic of 0.47 ($p = 0.82$) indicating that they jointly explain little of the variation in spending across states. These results suggest that there is little or no evidence that income distribution, physician specialty distribution, HMO penetration, medical malpractice insurance premiums, Medicaid participation and proportion of the population that is black, have an effect on aggregate state healthcare spending.

The data provide moderate to strong evidence that the variables $INC, EDU, URB, INS, DOC, BED, A_{65}, CIG, ALC, BMI$ and $EXC$ have independent effects on healthcare spending across states, with $t$-statistics that range from 1.61 ($p = 0.12$) for exercise to 3.74 ($p < 0.01$) for income. The algebraic signs of the estimates indicate higher healthcare spending for states with higher levels of income, less educated populations, fewer uninsured residents, larger elderly populations, greater availability of medical care providers, less urbanization and populations with less healthy lifestyles. In terms of size of effects, the magnitudes of elasticity estimates indicate that education has the biggest effect on healthcare spending across states (−0.51), followed by insurance coverage (0.43), income (0.37), obesity (0.21), availability of physicians (0.16), elderly population (0.12), urbanization (−0.10), availability of hospital beds (0.09), alcohol consumption (0.09), exercise (−0.08) and cigarette consumption (0.05).

Structural healthcare spending equation and health production function

Columns 1 and 2 of Table 2 present the results of 3SLS estimation for the health production function and structural spending equation. Estimates of structural coefficients provide information on direct and indirect (health) effects of factors that influence healthcare spending, and therefore illuminate the pathways through which they operate.

To assess the validity of the instrument sets, we test the overidentifying restrictions for the two structural Equations 5 and 6, and the joint significance of the
identifying instruments in the first-stage regression equation. For the spending equation (column 2), an approximate F-statistic of 1.13 ($p = 0.38$) provides strong evidence that the overidentifying restrictions are supported by the data, and therefore the instruments are exogenous. In addition, we resoundingly reject the null hypothesis that the identifying instruments have no joint effect on health status in the first-stage regression with an F-statistic of 28.29 ($p < 0.001$). The size of the F-statistic indicates that the instruments are strong. Together, these tests support the validity of our instruments for the structural spending equation and provide evidence that the elderly population, and the lifestyle related factors alcohol, cigarette consumption, exercise, and obesity do not have a direct effect on healthcare spending; rather they influence spending indirectly by affecting the health status of a state’s population.

The test statistics for the health production function (column 1) are less convincing and cast doubt on the validity of the instrument set used to identify this equation. An F-statistic of 35.36 ($p < 0.001$) strongly rejects the overidentifying restrictions indicating that one or more instruments are likely endogenous. While the F-statistic of 4.25 ($p = 0.003$) for the joint significance of the identifying instruments in the first-stage regression indicates significance at conventional levels, it falls well below the magnitude of 10 recommended by Stock and Watson (2003) suggesting that the variables $INS$, $MED$, $DOC$, $BED$, $HMO$ and $MAL$ are weak instruments, and therefore may produce estimates with substantial bias. Because the instruments appear to be relatively weak and one or more endogenous, caution should be exercised when interpreting the parameter estimates for the health production function and using these estimates to make inferences about indirect health effects.

Parameter estimates for the health production function, which are elasticities are given in column 1 of Table 2. These parameters along with the coefficient of the health status variable in the structural spending equation capture indirect health effects, which is a potentially important pathway through which socio-economic, lifestyle and various other factors may influence healthcare spending. The estimated coefficient of healthcare spending of $-0.15$ with a t-statistic of 0.84 is consistent with prior empirical studies of the aggregate health production function that find a 1% increase in medical care expenditures lowers mortality by approximately 0.05–0.15%, with most past estimates statistically
insignificant (Auster et al., 1969; Hadley, 1982; Thornton, 2002). The lifestyle related factors, cigarette and alcohol consumption, obesity and exercise have expected signs and are significant at the 5 and 1% levels, with the exception of alcohol consumption, which is insignificant. Urbanized states have significantly lower mortality, while states with larger elderly and black populations have significantly higher death rates. For socioeconomic variables, the results indicate that states with higher levels of income, less educated populations and greater disparity in income distribution have lower mortality rates. While the estimates for education and income distribution are unexpected and inconsistent with past studies, all three socio-economic variables are highly insignificant with t-statistics of 0.85 or less indicating the coefficients are imprecisely measured and providing no evidence of health effects, and therefore indirect spending effects.

Coefficient estimates measured as elasticities for the structural healthcare spending equation are provided in column 2 of Table 2. Recall that these structural parameters measure the direct effect of a change in a factor on healthcare spending, independent of any indirect effects on spending through population health status produced by the factor, as well as effects on spending that result from feedback interaction between healthcare spending, utilization and health status. The data provide relatively strong evidence that health status, income, education and availability of hospital beds have direct effects on healthcare spending. The t-statistics for the estimated effects of these variables range from 2.09 for education to 4.66 for income, with p-values of 0.04 and less. A 10% decrease in a state’s education or a 10% increase in income, mortality rate and hospital beds directly increases healthcare spending by 5.63, 4.70, 3.11 and 1.14%, respectively. Evidence of a direct effect is much weaker for urbanization, insurance, availability of physicians and income distribution with t-statistics between 1.13 and 1.26. For these variables, income distribution has the largest effect (0.35) followed by insurance (0.29), availability of physicians (0.07) and urbanization (−0.06); however, these effects are imprecisely measured. Finally, there is little or no evidence of a direct effect for specialty distribution, black population, HMO penetration, malpractice insurance premiums, and Medicaid, with t-statistics of 0.83 and less.

Potentially weak instruments, questionable identification and insignificant estimates with unexpected signs for several key variables suggest that the health production function may be misspecified, therefore, producing unreliable 3SLS estimates of structural parameters. An alternative approach to assess the pathways through which factors affect healthcare spending is to estimate the structural spending equation separately using 2SLS to obtain estimates of direct effects. The difference between estimates of total effects obtained from the reduced-form spending equation provided in Table 1 and direct effects from the 2SLS structural spending equation given in column 3 of Table 2 yield approximate estimates of indirect effects. The smaller the feedback effects the better the approximation. Column 3 of Table 2 presents the 2SLS estimates of the parameters of the structural spending equation. The 2SLS and 3SLS coefficient estimates are very similar indicating that estimates of direct effects are robust to these two alternative methods of estimation. Like the reduced-form spending equation, coefficient estimates of DIST, HMO, SPC, MED, MAL and BLK are highly insignificant with t-statistics of one or less. An approximate F-statistic of 0.29 (p = 0.94) indicates no joint direct effects for these variables.

A comparison of the estimated total effect (0.37) and direct effect (0.46) for income suggests that, as expected, indirect health and feedback effects are negative, and offset about 20% of the direct spending effect. For education, the estimate of the direct effect of −0.52 is almost identical to the estimate of the total effect of −0.51, providing no evidence of indirect health effects. For the other healthcare spending factors, imprecision of the estimates of direct and/or total effects provides no compelling evidence of indirect health effects for these variables, since differences in estimates of total and direct effects is likely due to sampling error.

Parsimonious specifications and robustness

The above results provide no convincing evidence that income distribution, physician specialty distribution, HMO penetration, medical malpractice insurance premiums, Medicaid participation, and black population have an effect on aggregate state healthcare spending. A more parsimonious specification for the OLS reduced-form spending equation and the 2SLS structural spending equation, which exclude these six variables, is given in columns 1 and 2 of Table 3.6

6 For the structural spending equation, an approximate F-statistic of 1.95 (p = 0.12) does not reject the overidentifying restrictions at conventional significance levels. An F-statistic of 24.41 (p < 0.001) for the joint significance of the identifying instruments in the first-state regression indicates strong instruments.
Table 3. Restricted reduced-form and structural healthcare spending equations (t-statistics in parentheses)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Reduced-form spending equation (OLS)</th>
<th>Structural spending equation (2SLS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>3.630 (4.31)</td>
<td>2.168 (2.23)</td>
</tr>
<tr>
<td>CDR</td>
<td>0.376 (4.15)</td>
<td>0.444 (4.58)</td>
</tr>
<tr>
<td>INC</td>
<td>−0.563 (2.53)</td>
<td>−0.430 (2.08)</td>
</tr>
<tr>
<td>EDU</td>
<td>−0.850 (2.13)</td>
<td>−0.065 (1.47)</td>
</tr>
<tr>
<td>INS</td>
<td>0.274 (1.48)</td>
<td>0.020 (0.10)</td>
</tr>
<tr>
<td>DOC</td>
<td>0.138 (3.22)</td>
<td>0.081 (2.06)</td>
</tr>
<tr>
<td>BED</td>
<td>0.114 (2.84)</td>
<td>0.110 (2.51)</td>
</tr>
<tr>
<td>A65</td>
<td>0.157 (3.42)</td>
<td></td>
</tr>
<tr>
<td>CIG</td>
<td>0.046 (2.61)</td>
<td></td>
</tr>
<tr>
<td>ALC</td>
<td>0.093 (1.98)</td>
<td></td>
</tr>
<tr>
<td>BMI</td>
<td>0.212 (3.19)</td>
<td></td>
</tr>
<tr>
<td>EXC</td>
<td>−0.083 (1.88)</td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.886</td>
<td></td>
</tr>
<tr>
<td>Test statistics (p-values in parentheses)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overidentifying restrictions–F-test</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No joint effect of identifying instruments–F-test</td>
<td>24.41 (&lt;0.001)</td>
<td></td>
</tr>
</tbody>
</table>

The reduced-form $R^2$ statistic of 0.886 indicates little difference in fit, relative to the more general specification in Table 1. Most coefficient estimates for the parsimonious equations are similar to those reported for the general specifications, but tend to have noticeably larger t-statistics indicating stronger evidence of effects. The exceptions are education and health insurance. The magnitude of the estimate of the total effect of education on healthcare spending, rises to −0.56, while the estimate of the direct effect falls to −0.43, with the former significant at the 1% level and the latter at the 5% level. These point estimates suggest that the direct effect accounts for about 77% of the total effect of education on healthcare spending, while indirect health and feedback effects account for the remaining 23%. The size of the estimate of the total effect of health insurance on spending decreases from 0.43 to 0.27, and is less precisely measured.

We find relatively strong evidence of higher levels of healthcare spending for state populations with higher income, less education, fewer uninsured residents, more elderly residents, less healthy lifestyles, greater availability of healthcare providers and less urbanization. Our results also help to illuminate the pathways through which these factors influence spending, and suggest an important role for population health status and its determinants. Lifestyle related factors, such as alcohol and tobacco consumption, exercise and obesity, as well as age and degree of urbanization exert their influence on medical care spending indirectly by affecting the health of a state’s population. Alternatively, there is some evidence that income may affect expenditures through both direct and indirect channels, but the direct pathway is clearly dominant. An important result is that education has the biggest effect on healthcare expenditures across states, and does so by directly affecting spending and possibly through its influence on population health status, but evidence for the latter channel is much weaker. We find little or no evidence of an effect on healthcare expenditures for the supply-side factors HMO penetration, medical malpractice

VI. Discussion and Conclusions

Column 1 of Table 4 reports the OLS estimates for the parsimonious specification of the structural spending equation. As expected, the OLS estimate of $CDR$ of 0.26 is less than the 2SLS estimate of 0.38 reflecting downward bias from reverse causation of healthcare spending on health status. A Hausman test of exogeneity of $CDR$ indicates that the bias is significant at the 1% level. The number of physicians per capita may also be an endogenous variable since it is possible that physicians tend to locate in states with higher healthcare spending. To examine this possibility, we re-estimate the structural spending equation treating both $CDR$ and $DOC$ as endogenous variables (column 2, Table 4). We also estimate a quasi reduced-form spending equation treating $DOC$ as an endogenous variable (column 3, Table 4). However, we find no compelling evidence that $DOC$ is endogenous in either the structural or reduced-form spending equations.7

7 We choose as instruments for $DOC$, the number of medical schools in a state and population density. For the quasi reduced-form equation, an approximate F-test of the overidentifying instruments provides evidence of exogeneity ($F = 0.07; p = 0.80$), and an F-test of their joint significance in the first-stage regression ($F = 9.49; p < 0.01$) provides evidence of their relevance. A Hausman test of exogeneity of $DOC$ cannot be rejected at any reasonable level of significance ($t = 1.18; p = 0.24$). For the structural spending equation, we also cannot reject the overidentifying restrictions ($F = 1.15; p = 0.35$). A general methods of moments test (Newey, 1985) of the hypothesis that $DOC$ is exogenous assuming $CDR$ is endogenous cannot be rejected at the 5% level of significance ($t = 1.88; p = 0.07$).
insurance cost and physician specialty distribution and the demand-side factors income distribution, Medicaid participation and African-American population. The factor that has arguably received the most attention in the empirical literature on healthcare spending is income, with an ongoing debate about whether healthcare is a luxury good or a necessity. Like prior studies that have analysed aggregate relationships, we find that income is a statistically significant and quantitatively important factor affecting healthcare spending with an income elasticity estimate of 0.38 (total effect), which is robust to alternative specifications. This estimate is smaller in magnitude than those reported at the national level, which tend to exceed 1.0, and less than half as large as Freeman’s (2003) aggregate state level estimates of 0.82–0.84. Freeman’s (2003) larger income elasticity estimate may reflect omitted variables that vary across states and over time for which he includes no explanatory control variables, and his healthcare spending measure, which does not adjust for border-crossing or state price level differences. The size of our estimates are consistent with Getzen’s (2000) multilevel model of healthcare spending that implies healthcare is a luxury good at the national level due to a binding country level resource constraint, but a necessity at lower levels of aggregation, such as the state, regional, county and individual levels, where budget constraints are relaxed by health insurance and government financing.

Getzen’s (2000) model also implies that income should have a substantially larger effect than population health status on healthcare spending at the state level of aggregation; however, this implication is inconsistent with our results. Our estimate of the health status elasticity of expenditures of 0.38 from the parsimonious specification of the structural spending equation is strongly significant and relatively close in size to the income elasticity estimate of 0.44 (direct effect). In addition, the sum of the magnitudes of the elasticities of the health status determinants $A_{65}$, $CIG$, $ALC$, $BMI$ and $EXC$ of 0.59 is greater than the income elasticity estimate of 0.38 (total effect). The individual health determinants with the largest estimated impact on spending are $BMI$ and $AGE$ with elasticities of 0.21 and 0.16. This suggests that increase in the proportion of the obese population and elderly have a similar impact on healthcare spending. Moreover, the effect of obesity on healthcare spending is about 2–5 times larger than

---

**Table 4. OLS structural healthcare spending equation and 2SLS quasi reduced-form and structural healthcare spending equations (physicians endogenous) ($t$-statistics in parentheses)**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Structural spending equation, OLS</th>
<th>Structural spending equation $CDR$, $DOC$ endogenous</th>
<th>Quasi reduced-form spending equation, $DOC$ endogenous</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>2.568 (2.52)</td>
<td>1.865 (1.87)</td>
<td>3.500 (4.57)</td>
</tr>
<tr>
<td>$CDR$</td>
<td>0.263 (3.83)</td>
<td>0.346 (4.60)</td>
<td></td>
</tr>
<tr>
<td>$INC$</td>
<td>0.407 (4.01)</td>
<td>0.469 (4.76)</td>
<td>0.427 (4.71)</td>
</tr>
<tr>
<td>$EDU$</td>
<td>-0.506 (2.35)</td>
<td>-0.500 (2.38)</td>
<td>-0.727 (3.07)</td>
</tr>
<tr>
<td>$URB$</td>
<td>-0.042 (0.91)</td>
<td>-0.031 (0.63)</td>
<td>-0.059 (1.41)</td>
</tr>
<tr>
<td>$INS$</td>
<td>0.200 (0.95)</td>
<td>0.164 (0.74)</td>
<td>0.421 (2.09)</td>
</tr>
<tr>
<td>$DOC$</td>
<td>0.072 (1.75)</td>
<td>0.022 (0.43)</td>
<td>0.068 (1.02)</td>
</tr>
<tr>
<td>$BED$</td>
<td>0.154 (3.51)</td>
<td>0.131 (2.91)</td>
<td>0.131 (3.41)</td>
</tr>
<tr>
<td>$A_{65}$</td>
<td></td>
<td></td>
<td>0.153 (3.69)</td>
</tr>
<tr>
<td>$CIG$</td>
<td></td>
<td></td>
<td>0.052 (3.13)</td>
</tr>
<tr>
<td>$ALC$</td>
<td></td>
<td></td>
<td>0.067 (1.44)</td>
</tr>
<tr>
<td>$BMI$</td>
<td></td>
<td></td>
<td>0.146 (1.85)</td>
</tr>
<tr>
<td>$EXC$</td>
<td></td>
<td></td>
<td>-0.063 (1.46)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.808</td>
<td>0.799</td>
<td>0.875</td>
</tr>
<tr>
<td>Test statistics ($p$-values in parentheses)</td>
<td>1.15 (0.35)</td>
<td>0.07 (0.80)</td>
<td>9.49 (&lt;0.001)</td>
</tr>
<tr>
<td>Overidentifying Restrictions–$F$-test</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No joint effect of identifying restrictions–$F$-test</td>
<td></td>
<td></td>
<td>3.40 (&lt;0.01)</td>
</tr>
<tr>
<td>Exogeneity of $CDR$–$t$-test</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exogeneity of $DOC$–$t$-test</td>
<td></td>
<td></td>
<td>1.18 (0.24)</td>
</tr>
<tr>
<td>Exogeneity of $DOC$ assuming $CDR$ endogenous–$t$-test</td>
<td></td>
<td>1.88 (0.07)</td>
<td></td>
</tr>
</tbody>
</table>
alcohol and cigarette consumption. Our elasticity estimate for cigarette consumption of 0.05 is smaller than the estimate of about 0.13 reported in cross-country comparisons, but consistent with other studies that find little impact on healthcare spending over the life-cycle (Gertham and Jonsson, 2000).

The story that emerges from our results is that the income level and health status of a state’s population are equally important determinants of healthcare spending. Moreover, comparison of our total and direct effect income elasticity estimates suggest that at a maximum only about 14–20% of the direct effect of income on healthcare spending is offset by nonmedical care, income related improvements in health, while our estimate obtained from the health production function provides no evidence of an indirect income effect. This is consistent with empirical studies that analyse the effect of income on health, which tend to find a relatively large effect at the individual level, but a negligible effect at the aggregate level (Deaton, 2002). Our finding that population health status has an important effect on healthcare spending also helps to explain the frequently observed negative correlation between medical care spending and measures of health at the state, regional and national level. Aggregate level cross-section studies of health outcomes in the US and other developed countries suggest that the aggregate marginal effect of medical care on health is relatively small and statistically insignificant indicating ‘flat-of-the-curve’ medicine (Auster et al., 1969; Fuchs and Kramer, 1980; Thornton, 2002). The findings of these studies together with our results suggest that causation runs primarily from health to medical care, rather than from medical care to health.

Our results suggest that demand-side factors are relatively more important than supply-side factors in explaining variation in healthcare spending across states. In our most general models, of the five factors that potentially influence expenditures through their influence on supply and cost of providing care, compelling evidence of an effect exists for only two, physician and hospital availability, and their influence may well work through a demand mechanism, rather than or in addition to, a supply mechanism. Many have argued that the medical malpractice environment is a major factor driving up healthcare spending, and HMOs are an effective delivery system for reducing costs and expenditures. We find no evidence in the aggregate state level data to support these assertions. It may be that in the managed care environment of the latter 1990s physicians were unable to pass along premium costs to patient fees, and therefore absorbed most of these costs through reductions in net incomes. What is more, variations in spending that are said to reflect defensive medicine may be insurance-induced and explained by variation in insurance coverage (Danzon, 1991). However, our measure of liability premiums is admittedly subject to measurement error, since we are required to use census division average premiums as a proxy for 40 states, and therefore our estimate of the effect of malpractice cost may be biased towards zero. Similarly, our inability to detect an effect of HMO penetration in the aggregate data may be the result of bias from reverse causation; that is, HMO health plans may be more likely to arise in states with higher healthcare spending in an effort to contain cost. However, while prior research finds that HMO health plans have the ability to reduce cost, evidence from empirical studies remains inconclusive about the effect on aggregate healthcare spending (Glied, 2000).

Like many prior studies, we find that geographic regions with more physicians and hospital beds per capita have higher healthcare spending. Several alternative explanations are consistent with this result. States with higher healthcare spending may have greater demand for physician and hospital services, which attracts more doctors and elicits increased bed supply from hospitals. However, we find no compelling evidence that DOC is endogenous, and therefore we have no reason to believe that causation runs from healthcare spending to physician availability. In addition, Feldstein (1971) argues there is no convincing evidence that hospital bed supply in the US reflects demand patterns. Higher spending associated with greater availability of providers may also reflect supplier-induced demand and provision of unnecessary medical care or lower waiting and travel times that increase demand for both physician and hospital services (Feldstein, 1971; Frech, 1996). To conclude, our results suggest that greater availability of physicians and hospital beds leads to higher state healthcare spending, but we can only speculate about the channel(s) through which this effect operates.

The potential role of education in aggregate healthcare spending has been overlooked in previous research. To our knowledge, no prior studies have

---

8 As Paul Feldstein has observed, if physicians were able to pass on higher malpractice insurance costs to patients and third-party payers, then “it is unlikely they would spend their time in marches on their respective state capitals seeking legislative relief” (1988, p. 188).

9 We estimated our general models using several alternative measures of MAL. (1) Premium measure reported in the article not adjusted for state cost of living differences. (2) Premium measure using census division average premiums as a proxy for all 50 states, with and without cost of living adjustments. Our results were very similar for all of these alternative measures.
analysed the aggregate relationship between education and healthcare spending. Our results suggest that the education level of a state’s population is as important as its level of income in determining how much it spends on healthcare services, and may be the most important factor explaining variation in expenditures across states. The elasticity estimate of -0.56 (column 1, Table 3) indicates that a 10% increase in the proportion of the adult population with a high school education leads to a 5.6% reduction in spending on healthcare services. Somewhat surprisingly, our results suggest that education influences spending primarily through a direct effect that operates independent of its effect on health. A direct negative effect is inconsistent with several micro level studies discussed previously that find individuals with more years of schooling tend to spend more on certain types of medical care services. A possible explanation for this discrepancy involves the existence of state group-level education effects that dominate individual effects. It may be that within states, more educated individuals tend to spend more on medical care, but the average level of healthcare spending is lower in states that have a higher average level of education. This may result from the existence of group-level education externality effects that increase the average efficiency with which individuals in the states with higher average levels of education purchase and utilize medical care.

The results of our study have potentially interesting and important policy implications. In response to rapidly rising healthcare spending and attendant budgetary and medical care access problems, states are taking actions to contain healthcare costs. Many state legislatures have proposed or implemented supply oriented reform initiatives that involve managed care, medical malpractice tort reforms emphasizing caps on damage awards and restrictive certificate of need (CON) laws that regulate expansion of healthcare facilities and new medical equipment. We find no compelling evidence that actions to curb medical malpractice insurance cost and foster growth of HMO will be effective at containing state healthcare spending. Our analysis does suggest that CON laws designed to limit availability of hospital beds may be effective in curbing healthcare spending; however, historically these laws have proven costly to administer, and therefore may not be a cost-effective policy action.

Our findings suggest that the most effective cost containment measures may be those that promote healthy lifestyles and improve educational opportunities. These types of policy actions may be particularly cost-effective because they lead to reductions in healthcare spending and help to achieve other potential social policy goals as well, such as improving population health status and increasing the productivity of the work force. Recent research suggests that social benefits of education may also include lower levels of criminal activity and greater participation in the political process (Lochner and Moretti, 2004; Dee, 2004). Moreover, studies find that education is negatively correlated with behavioural risk factors such as cigarette consumption, excessive use of alcohol, sedentary lifestyle and obesity. To the extent that more formal education leads to healthier lifestyle choices, this will result in additional improvements in health and further reductions in spending over and above those produced by the independent effect of education.

References
Determinants of healthcare spending


Appendix: Analytical Framework

Our analytical framework consists of Cobb–Douglas type structural equations for the quantity of healthcare \( Q \), price of healthcare \( P \) and health status \( H \) and a healthcare spending identity \( S \), given by

\[
Q = \alpha_0 X^{\alpha_X} Z^{\alpha_Z} H^{\alpha_H} \quad (A1)
\]

\[
P = \phi_0 X^{\phi_X} Z^{\phi_Z} P^{\phi_P} \quad (A2)
\]

\[
H = \beta_0 Y^{\beta_Y} Z^{\beta_Z} Q^{\beta_Q} \quad (A3)
\]

\[
S = P Q \quad (A4)
\]

where \( X \) are exogenous variables that have a direct effect on spending, \( Y \) are exogenous variables that have an indirect effect on spending by affecting health status, \( Z \) are exogenous variables that have both direct and indirect effects on spending and \( \alpha_0, \alpha_X, \alpha_Z, \alpha_H, \phi_0, \phi_X, \phi_Z, \phi_H, \beta_0, \beta_Y, \beta_Z \) and \( \beta_Q \) are parameters. Exogenous variables in \( X \) and \( Z \) also fall into two alternative categories: demand-side variables designated \( X^d \) and \( Z^d \) and supply-side variables denoted \( X^s \) and \( Z^s \). We expect \( X^d \) and \( Z^d \) factors that increase (decrease) demand to have a positive (negative) effect on price and quantity of healthcare, and therefore \( \alpha_X^d, \alpha_Z^d, \phi_X^d, \phi_Z^d > 0 \). We expect \( X^s \) and \( Z^s \) factors that increase (decrease) supply to have a negative (positive) effect on price and positive (negative) effect on quantity, and therefore \( \alpha_X^s, \alpha_Z^s > 0 \) and \( \phi_X^s, \phi_Z^s < 0 \). We allow for factors that may affect both supply and demand. Lastly, we expect an improvement (decline) in population health status to decrease (increase) demand, price and quantity of healthcare services, and therefore \( \beta_H < 0, \beta_Z < 0 \).

Substituting (A1) and (A2) into (A4) and simplifying yields the following structural healthcare spending equation

\[
S = \gamma_0 X^{\gamma_X} Z^{\gamma_Z} H^{\gamma_H} \quad (A5)
\]

where \( \gamma_X = \alpha_X + \phi_X \); \( \gamma_Z = \alpha_Z + \phi_Z \); \( \gamma_H = \alpha_H + \phi_H \). Substituting (A1) into (A3) and solving for \( H \), and then substituting the resulting expression for \( H \) into (A5) and simplifying yields the following reduced-form healthcare spending equation

\[
S = \pi'_0 X^{\pi_X} Y^{\pi_Y} Z^{\pi_Z} \quad (A6)
\]

The reduced-form slope coefficients are given by the following expressions

\[
\pi_X = \gamma_X + \left( \frac{\alpha_X \beta_H \gamma_H}{\lambda} \right) \quad (A7a)
\]

\[
\pi_Y = \frac{\beta_Y \gamma_H}{\lambda} \quad (A7b)
\]

\[
\pi_Z = \gamma_Z + \left( \frac{\alpha_Z \beta_Q \gamma_H}{\lambda} \right) + \left( \frac{\beta_Z \gamma_H}{\gamma} \right) \quad (A7c)
\]

where \( \lambda = 1 - \alpha_H \beta_Q \). Taking natural logarithms of the reduced-form and structural healthcare spending Equations A6 and A5 yields the same Equations 1 and 2

\[
\ln S = \pi_0 + \pi_X \ln X + \pi_Y \ln Y + \pi_Z \ln Z \quad (A8)
\]

\[
\ln S = \gamma_0 + \gamma_X \ln X + \gamma_Y \ln Y + \gamma_Z \ln Z + \gamma_H \ln H \quad (A9)
\]

The coefficients of the structural spending Equation 2, \( \gamma_i \), provide information on the direct effects of \( X \), \( Z \) and \( H \) on healthcare spending. We expect \( X^d \) and \( Z^d \) factors that increase (decrease) demand to have a positive (negative) effect on healthcare spending, and therefore \( \gamma_X^d, \gamma_Z^d > 0 \). Assuming an inelastic demand for healthcare services, which is consistent with prior empirical studies, we expect \( X^s \) and \( Z^s \) factors that increase (decrease) supply, to have a negative (positive) effect on healthcare spending, and therefore \( \gamma_X^s, \gamma_Z^s < 0 \).

The coefficients of the reduced-form spending Equation 1, \( \pi_i \), measure the total (marginal) effect of a change in a factor that influences healthcare spending, and can be decomposed into three
possible effects. (1) Direct effect of a factor on spending. (2) Indirect effect of a factor on spending through health status. (3) Feedback effects on spending that result from the spending-health interaction. Variables in $X$ have a direct effect on spending measured by $\gamma_{X}$; variables in $Y$ have an indirect effect (via health status) on spending given by $\beta_{Y|H}$; and variables in $Z$ have both direct and indirect effects on spending given by $\gamma_{Z}$ and $\beta_{Z|H}$. All three types of variables produce feedback effects. The feedback effects produced by a direct effect is given by $\frac{\alpha_{i} \beta_{Q|H}}{\lambda}$ for $i = X, Z$, while those for an indirect effect are captured by $(1/\lambda)$.

To analyse factors that influence healthcare spending across states and the pathways through which they operate (total, direct and indirect effects), Equations 1 and 2 and the logarithmic form of the health production function, A3 can be estimated, provided that $S$ is used as a proxy for unobservable $Q$ in A3. Our framework suggests that an alternative strategy would be to estimate the structural Equations A1, A2 and A3 directly, and use the resulting parameter estimates to draw inferences about factors influencing healthcare spending and the mechanisms involved. However, this is not a viable strategy because $P$ and $Q$ are not directly observable, and state healthcare price indices do not exist to obtain an estimate of the quantity of real healthcare services by state.

### Table A1. Variable definitions and descriptive statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
<th>Mean</th>
<th>SD</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S$</td>
<td>Healthcare expenditures per capita adjusted for price level</td>
<td>3748.36</td>
<td>379.97</td>
<td>2733.73</td>
<td>4443.96</td>
</tr>
<tr>
<td>$CDR$</td>
<td>Death rate/100 000 population</td>
<td>873.4</td>
<td>133.07</td>
<td>420.0</td>
<td>1150.0</td>
</tr>
<tr>
<td>$BMI$</td>
<td>% Of population age 18+ with body mass index of 30+</td>
<td>18.06</td>
<td>2.51</td>
<td>13.1</td>
<td>23.9</td>
</tr>
<tr>
<td>$ALC$</td>
<td>Gallons of wine, spirits and beer per capita</td>
<td>2.26</td>
<td>0.472</td>
<td>1.25</td>
<td>4.04</td>
</tr>
<tr>
<td>$CIG$</td>
<td>Packs sold per capita</td>
<td>45.62</td>
<td>14.97</td>
<td>9.61</td>
<td>103.6</td>
</tr>
<tr>
<td>$EXC$</td>
<td>% Of population age 18+ who exercise 5+ times/30 minutes per week</td>
<td>20.48</td>
<td>4.07</td>
<td>13.0</td>
<td>30.0</td>
</tr>
<tr>
<td>$A65$</td>
<td>% Of population age 65+</td>
<td>12.71</td>
<td>1.96</td>
<td>5.50</td>
<td>18.30</td>
</tr>
<tr>
<td>$INC$</td>
<td>Personal income per capita adjusted for price level</td>
<td>25,373.23</td>
<td>2926.66</td>
<td>20,468.17</td>
<td>33,450.18</td>
</tr>
<tr>
<td>$DIST$</td>
<td>Gini coefficient of income inequality</td>
<td>0.4277</td>
<td>0.0219</td>
<td>0.385</td>
<td>0.476</td>
</tr>
<tr>
<td>$EDU$</td>
<td>% Of population age 25+ who have at least high school education</td>
<td>83.97</td>
<td>4.45</td>
<td>76.4</td>
<td>92.0</td>
</tr>
<tr>
<td>$INS$</td>
<td>% Of population with insurance</td>
<td>84.98</td>
<td>3.97</td>
<td>75.5</td>
<td>91.0</td>
</tr>
<tr>
<td>$MED$</td>
<td>% Of population with Medicaid</td>
<td>13.58</td>
<td>4.87</td>
<td>4.63</td>
<td>33.95</td>
</tr>
<tr>
<td>$DOC$</td>
<td>Physicians per 100 000 population</td>
<td>234.34</td>
<td>57.62</td>
<td>154.0</td>
<td>412.0</td>
</tr>
<tr>
<td>$BED$</td>
<td>Hospital beds per 100 000 population</td>
<td>326.62</td>
<td>100.82</td>
<td>189.0</td>
<td>624.0</td>
</tr>
<tr>
<td>$HMO$</td>
<td>% Of population not enrolled in an HMO</td>
<td>76.82</td>
<td>14.31</td>
<td>45.8</td>
<td>100.0</td>
</tr>
<tr>
<td>$MAL$</td>
<td>Average liability premiums for self-employed physicians (thousands of $) adjusted for price level</td>
<td>15.53</td>
<td>3.39</td>
<td>8.20</td>
<td>21.65</td>
</tr>
<tr>
<td>$SPC$</td>
<td>% Of nonprimary care physicians</td>
<td>72.46</td>
<td>5.33</td>
<td>59.9</td>
<td>81.7</td>
</tr>
<tr>
<td>$URB$</td>
<td>% Of population living in MSA: Metropolitan Statistical Area</td>
<td>67.71</td>
<td>20.75</td>
<td>27.9</td>
<td>100.0</td>
</tr>
<tr>
<td>$BLK$</td>
<td>% Of population that is black</td>
<td>10.17</td>
<td>9.58</td>
<td>0.341</td>
<td>36.45</td>
</tr>
</tbody>
</table>
### Table A2. Data sources

<table>
<thead>
<tr>
<th>Variable</th>
<th>Data source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>S</strong></td>
<td>Healthcare spending per capita adjusted for border-crossing taken from (Martin et al., 2002). State cost of living index taken from Leonard et al. (1999).</td>
</tr>
<tr>
<td><strong>A65</strong></td>
<td>US Census Bureau. Data taken from <em>statistical abstract of US, 1999</em>, Table 33.</td>
</tr>
<tr>
<td><strong>BMI</strong></td>
<td>Centers for Disease Control and Prevention, Behavioural Risk Factor Surveillance System. Data obtained from <a href="http://www.cdc.gov">www.cdc.gov</a></td>
</tr>
<tr>
<td><strong>EXC</strong></td>
<td>Centers for Disease Control and Prevention, Behavioural Risk Factor Surveillance System. Data obtained from <a href="http://www.cancer.org">www.cancer.org</a></td>
</tr>
<tr>
<td><strong>CIG</strong></td>
<td>Annual gross tax revenues and federal and state cigarette tax per pack from Centers for Disease Control and Prevention, State Tobacco Activities Tracking System. Data obtained from <a href="http://www.cdc.gov">www.cdc.gov</a></td>
</tr>
<tr>
<td><strong>ALC</strong></td>
<td>National Institute on Alcohol Abuse and Alcoholism database. Data obtained from <a href="http://www.niaaa.nih.gov">www.niaaa.nih.gov</a></td>
</tr>
<tr>
<td><strong>HMO</strong></td>
<td>Interstudy. Data taken from <em>Health, United States, 2002</em>, Table 146.</td>
</tr>
<tr>
<td><strong>URB</strong></td>
<td>US Census Bureau. Data taken from <em>statistical abstract of US, 2000</em>, Table 33.</td>
</tr>
</tbody>
</table>