Health care is an individual necessity and a national luxury: applying multilevel decision models to the analysis of health care expenditures

Thomas E. Getzen *

Department of Risk, Insurance and Healthcare Management, Temple University, Speakman Hall, Philadelphia, PA 19122, USA
International Health Economics Association, USA

Abstract

Health care is neither “a necessity” or “a luxury”; it is “both” since the income elasticity varies with the level of analysis. With insurance, individual income elasticities are typically near zero, while national health expenditure elasticities are commonly greater than 1.0. The debate over whether health care is or is not a luxury good arises primarily from the failure to specify levels of analysis clearly so as to distinguish variation within groups from variation between groups. Apparently, contradictory empirical results are shown to be consistent with a simple nested multilevel model of health care spending. © 2000 Elsevier Science B.V. All rights reserved.

JEL classification: I1
Keywords: Income elasticity; Multilevel analysis; Health expenditures; Aggregation; Unit of observation

A debate has gone on for more than a decade in the pages of JHE and other journals as to whether or not health care is a luxury good, that is, whether the
income elasticity of expenditures is above or below 1.0 (Newhouse, 1987; Parkin et al., 1987; Gerdtham et al., 1992; Hitris and Posnett, 1992; Hansen and King, 1996; Blomqvist and Carter, 1997; Di Matteo and Di Matteo, 1998). Confusion and conflicting empirical results have arisen primarily from the failure to carefully specify the unit of observation. It is to be expected that measured income elasticities will differ for an individual, a risk-pooling group, or a national health system, just as price elasticities for individual, firm and market demand normally differ from each other. More concretely, the income elasticity of individual health expenditures under insurance (usually 60% to 95% of total spending) is typically near zero or negative, while the elasticity of national health expenditures with respect to national income is typically greater than 1. This paper will demonstrate how multilevel decision modeling quickly resolves the apparent discrepancies, and make evident the crucial role of social and private insurance in linking micro and macro analysis in health economics.

1. Variation within groups and variation between groups

The basic analytical problem arises from a failure to make the distinction between sources of variation within groups and sources of variation between groups (for useful commentaries, see Newhouse, 1977; Susser, 1994; Forni and Lippi, 1997; Diez-Roux, 1998). Within an insured group, the bulk of the health resources will be allocated to those individuals who are ill and stand to benefit from medical care. Individual budget constraints and ability to pay concerns are largely removed through pooled insurance financing. However, variation between groups which do not share access to the same insurance pool is largely determined by differences in aggregate group funding. The group mean will vary with the budget constraint, not with the amount of illness.

Confusion arises when the analyst fails to recognize that \( \frac{\overline{x}}{\overline{y}} \) (where the overline denotes the group mean of a variable) is a statement about the behavior of the entire group expressed in terms of averages, not a statement about (or estimate of) the behavior of the “average individual” (Kirman, 1992). Most of the time, \( \frac{\overline{x}}{\overline{y}} \neq \frac{x_i}{y_i} \); in particular, if personal expenditures are influenced by the average community or group per capita income, as well as individual income, then equality cannot be maintained. Usually, \( \frac{\overline{x}}{\overline{y}} \) is estimated from data on individuals within one or more groups, while \( \frac{\overline{x}}{\overline{y}} \) is estimated by making comparisons across groups. An econometric issue that has been dealt with at length by Theil (1954) and others (Green, 1964; Goldstein, 1995) is the extent to which the two estimates would be the same if there were no independent group

\[1 \text{For the specific limited cases in which equality would be maintained, see the discussions in Theil (1954), Blalock (1964) or Goldstein (1995). Note also that } \frac{y}{x} \text{ (or } \frac{X}{Y} \text{) is not a function since the values could, and in fact are likely to, depend upon the variance or skewness of } y, \text{ as well as the mean.}\]
effects. The consistency of estimates issue has usually been addressed under the rubric of the “problem of aggregation” or “homology”. More recently, the question of estimating separate individual and group level effects has been raised, and a body of literature has been developed by Blalock (1964), Hannan (1991), Bryk and Raudenbush (1992) and Goldstein (1995) developing maximum likelihood techniques for estimation of non-zero group effects that can parse out the relative contribution of individual and group means (or distribution) under a set of linear but flexible assumptions.

The contrast between the behavior of the average individual, and the behavior of the group mean, is well-illustrated by insurance. Define “pooling groups” as bounded sets of individuals that pool funds by paying insurance premiums ($p$) in order to make health expenditures ($x$). Aggregate values are indicated by capital letters, so that ($P$) and ($X$) are total group premiums and total group expenditures, respectively. For simplicity, start with the assumptions that each member pays the same premium, that all expenditures are made from the pooled premiums, that these premiums are fully expended during a single period by the members of the group, and that the relevant measures of individual income and expenditures are well-defined. These restrictions can be relaxed to show the effects of multiple and/or overlapping groups, joint costs, self-paid expenditures, administrative overhead, taxation, subsidies and inter-period transfers. If the pool consists of a single individual, then group mean and individual average behavior are, by construction, identical.

$$[ N = 1 ] \quad x_1 = x = X \equiv P = p = p_1$$

The lack of external subsidy or leakage is shown by ($X \equiv P$), the group expenditures = premiums constraint. With two members in the pool, some individual (but not external) cross-subsidy is likely. Indeed the primary purpose of insurance is to make sure that $d p_1 / d x_1 < 1$.

$$[ N = 2 ] \quad x_1 + x_2 = 2 x = X \equiv P = 2 p = p_1 + p_2$$

As the number of members becomes larger, the individual connection between expenditures and premiums becomes ever more tenuous, while for the group mean, that is, for the pool as a whole, expenditures and premiums must maintain a 1:1 correspondence.

$$[ N = n ] \quad x_1 + \sum_{i=2}^{n} x_i = nx = X \equiv P = np = p_1 + \sum_{i=2}^{n} p_i \quad n \to \infty,$$

$$d p_i / d x_i \to 0, \quad d x / d p = 1$$

Since the group is self-funding, (i.e., $X \equiv P$), any increase in expenditures on individual must be offset by a slight decrease in the expenditures on all other individuals, or a slight increase in the premiums of all group members. For a large pool, the per-member offset is so small as to be virtually unnoticeable at the individual level ($N$ large, $d x_i / d x_j \approx 0 = d p_i / d x_i$).

The “allocation” of total expenditures across individuals can be defined as a fraction $a_i = x_i / X$. The change in expenditures on individual, due to a change in
any variable \( y \) can thus be decomposed into the change in allocation and the change in total group expenditures from the pool (and a cross-term that becomes negligibly small as groups become large).

\[
d x_i / d y_i = d a_i / d y_i \times X + a_i \times d X / d y_i
\]  

Estimates made across individuals within a pooling group capture all of the variance within the first “allocative” term. Conversely, estimates of the group mean necessarily zero out the first term (since mean allocation is a constant \( \bar{a} = 1/n \) by construction), and only the changes in total expenditures between groups matter.\(^2\)

\[
d \bar{X}/d \bar{y} = 0 + 1/n \times d X / d \bar{y}
\]  

Estimates of the coefficients \( d x_i / d y_i \) and \( d \bar{X}/d \bar{y} \) may thus be quite different, even opposite in sign. This difference in the determinants of the group mean as compared to the individual (idiosyncratic or average) response suggests a multi-level modeling strategy that treats the determination of average expenditures \( \bar{X} \) and the allocation of individual expenditures \( x_i \) as separable processes.

2. A two-level allocative model

If a pooling group is sufficiently large, members will tend to neglect the total budget constraint in making individual decisions (since the amount of incremental costs paid by any one person, \( \Delta P/n \), becomes vanishingly small) and the decision of how much of the total budget to allocate to a specific individual based on small scale micro factors \( (s) \) may be almost entirely separate from the large scale \( (L) \) macro factors determining the total constraint. In such a case, one can write a particularly simple nested multilevel hypothesis:\(^3\)

\[
X = X( L) \quad \text{budget determination (Macro, Level I)} \tag{7a}
\]

\[
x_i = a(s_i) \times X \quad \text{allocation of budget to individuals (micro, level II)} \tag{7b}
\]

Multilevel models become more useful and efficient as group decision-making becomes more important than individual behavior in determining \( \bar{X} \). If the “groups” are randomly constituted aggregations with no organization and no

\(^2\)The \( \bar{X} = 1/n \) equation is an accounting identity, not a behavioral assumption. It must be true by construction that for a group of 16 members, the average allocation is \( 1/16 \), for 25 members \( 1/25 \), and in general \( 1/n \).

\(^3\)Note that the multilevel model presented here differs substantially from that used by Goldstein (1995), Rice and Jones (1997), Scott and Shiell (1997) and most others in that it is a nested model. The determination of total expenditures is made separately from the determination of individual expenditures. In the more typical general linear model, there are effects at both Level I and Level II which affect the total. In a nested model, the group mean is, by construction, independent of the individual level allocative variables.
power over finances (e.g., all persons whose social security number ends in 7, all persons wearing blue socks on a given day), only micro level (s) variables matter. Estimation of group means in such random collections reveals individual behavior only, and is homologous across levels (Blalock, 1964, pp. 97–114). If the unit of observation is a mixed ‘‘group’’ with both higher and lower order effects (inhabitants of a census tract, patients with disease X, admissions to Hospital Z) then a single equation that mixes micro and macro variables such as $x_i = x(s, L)$ performs quite well (Bryk and Raudenbush, 1992). Allocative models separating group and individual choices work best when effects are clearly separated, and when budget constraints apply to the group more than to the individual within the group.

Insurance pools are not only likely to display separation between group and individual behavior, they are designed to bring about such a separation. The purpose of insurance is to remove the individual budget constraint, and to reduce or eliminate the influence of cost of care on patients’ and physicians’ decisions of how much care to use. Thus, we do not expect to find significant income effects on the utilization of care among members of an HMO, and only modest income effects among members of an indemnity plan with 10% coinsurance. If persons are fully insured, correlations with measures of individual income provide no information about income effects per se (e.g., the effect of monetary budget constraints) but instead reflect the influence of other unmeasured variables (cost of time, family resources, education, preferences, planning horizons, etc.) that are correlated with an individual’s income.

![Multilevel model of determinants of health expenditures](image)

**Fig. 1.** Multilevel model of determinants of health expenditures. Without insurance (left) both income and health status affect expenditures at the micro level. At the macro level, income effects are still strong, but variation due to differences in health status is no longer visible. The short arrows indicate that income also affects health status directly at both the micro and macro levels. With insurance (right), the pooling of funds removes the income constraints at the micro level, and tends to strengthen the correlation of individual health status with expenditures. At the macro level, income effects still dominate.
Fig. 1 presents the bifurcated model diagrammatically. At the individual micro level, without organized medical financing, both health status \( (s_i) \) and income \( (y_i) \) affect expenditures, but most of the variance is due to differences in individual health status. At the macro level in an uninsured state, the large individual variation due to differences in health status is averaged out and disappears, so that income effects, while no bigger in magnitude, become more visible as the background noise is removed. With the development of insurance, an organized link is created between the micro and macro levels which further attenuates the individual income effects, and reinforces the health status effects at the individual level. However, the aggregate effect of individual incomes on the group mean is not reduced or eliminated, merely pooled and transferred to a higher level. The model as constructed determines total spending at the national level based largely on the average amount of national resources (per capita income \( \bar{y} \)), while the allocation of expenditures across individuals is based primarily on their need for care \( (s_i) \) rather than their contribution to total resources. Note that it is not entirely correct to speak about “individual spending” in such a model since the expenditures on individuals are based on group decisions, with little (financial) opportunity cost foregone at the individual level.  

3. Empirical results

Most recent studies of individual expenditures show that the majority of the variation in spending (50%–90%) is associated with individual differences in health status, while income elasticities are small or negative (see Table 1) (Newhouse and Phelps, 1976; AMA, 1978; Manning et al., 1987; Sunshine and Dicker, 1987; Wedig, 1988; Wagstaff et al., 1991; AHCPR, 1997). However, analysis of pre-1960 data where insurance is less prevalent and most payment is made out-of-pocket shows a much larger income elasticity (0.2–0.7) (Falk et al., 1933; Anderson et al., 1960; Weeks, 1961). Similarly, consumption of dentistry, plastic surgery, counselling, eyeglasses, topicals, and other types of care that are still less well-insured show income elasticities that are strongly positive, and sometime substantially exceed 1.0 (USPHS, 1960; Andersen and Benham, 1970; Silver, 1970; AMA, 1978; Scanlon, 1980; Sunshine and Dicker, 1987; AHCPR, 1997; Parker and Wong, 1997). At the macro level, studies of national expenditures consistently show income elasticities greater than 1.0, with 90 + % of cross-sectional and time series variation explainable by differences in per capita income, and differences in health status as having negligible effects (Abel-Smith, 1967; Kleiman, 1974; Newhouse, 1977; Maxwell, 1981; Leu, 1986; Culyer, 1988; Getzen, 1990; Getzen and Poullier, 1992; Schieber, 1990; Gerdtham et al., 1992). 

---

4 Similar considerations arise in the modeling of consumption across members in a household, joint products, or of individual periods across the life-cycle.
Reconciling these disparities requires a multilevel analysis. Insurance, although removing the cost considerations from individual decision-making, leaves the budget constraint still binding on the group as a whole. The difference in income elasticities between insured and uninsured care is consistent with this interpretation, as are the intermediate elasticity values for semi-pooling groups which are only partially bounded financially (regions, states, counties, provinces). Note that it is the dominance of income effects at the national level that establishes a particularly simple form of group budget constraint, and makes the multilevel “allocative” model presented above particularly useful in the analysis of health care costs.

4. Open and closed pools: taxes, inter-temporal transfers and joint costs

Many apparently closed funding pools such as HMOs, and private insurance companies and employer benefit trusts, are actually porous mixed intermediates, with substantial leakage due to taxation, transfers, cross-subsidization of catastrophic cases, and government guarantees. The only set of coefficients that consistently reflects income effects within closed pools are those which use the nation as the unit of observation. Even for the nation as a whole, financial boundaries can be flexed by transfers of funds across time. Borrowing, unanticipated inflation and other factors can ease or distort the budget constraint. In order to truly capture the effects of income on health care spending, it is necessary to specify a distributed lag or ARIMA model spanning several years. The budget constraint is much more tightly binding in the long-run, and national time series estimates of income elasticity that account for lags are greater in magnitude and explanatory power, and also reveal that the correlation with current income is close to 0. This is not surprising since most elements of national expenditure must be budgeted a year or more in advance (Getzen, 1990; Getzen and Poullier, 1992).

Inter-temporal transfers apply at the individual level as well. A slight shift in notation and the incorporation of an interest rate would convert Eq. (2) into a standard two-period consumption model. Since payments are fungible between periods, consumption in each period depends more on total earnings overall rather than measured current earnings, results usually known as the permanent income life-cycle hypothesis (Friedman, 1957). The important point is that insurance and savings are both methods of pooling (Getzen, 1997, pp. 384–392) and pose econometric problems that are structurally similar: corr$(x, y)$ is biased toward 0 while the constant (which theoretically ought to be 0, and is 0 in long time series

---

5 Most nations are closed pooling groups with respect to national health expenditures. However, some developing countries obtain a substantial proportion (up to 50%) of total health budgets in the form of aid and transfers from international bodies (World Bank, 1993, Table A.9).
Table 1
Estimated income elasticity of health care expenditures (or utilization) by level of observation from a
variety of studies over the last 50 years. Since methodologies vary and elasticities must be interpolated
in many cases, readers are cautioned to carefully refer to original sources in the list of references.

<table>
<thead>
<tr>
<th>INDIVIDUALS [micro]</th>
<th>Income Elasticity</th>
</tr>
</thead>
<tbody>
<tr>
<td>General (insured / mixed)</td>
<td></td>
</tr>
<tr>
<td>Newhouse and Phelps (1976)</td>
<td>≤ 0.1</td>
</tr>
<tr>
<td>AMA (1978)</td>
<td>= 0</td>
</tr>
<tr>
<td>Sunshine and Dicker (1987) (NMCUES)</td>
<td>= 0</td>
</tr>
<tr>
<td>Manning et al. (1987) (Rand)</td>
<td>= 0</td>
</tr>
<tr>
<td>Wedig (1988) (NMCUES)</td>
<td>= 0</td>
</tr>
<tr>
<td>Wagstaff et al. (1991)</td>
<td>≤ 0</td>
</tr>
<tr>
<td>Hahn and Lefkowitz (1992) (NMES)</td>
<td>≤ 0</td>
</tr>
<tr>
<td>AHCPR (1997) (NMES)</td>
<td>≤ 0</td>
</tr>
<tr>
<td>Special / uninsured</td>
<td></td>
</tr>
<tr>
<td>Pre-1960 Expenditure Data</td>
<td></td>
</tr>
<tr>
<td>Falk et al. (1933)</td>
<td>0.7</td>
</tr>
<tr>
<td>Weeks, 1961 (1955 data)</td>
<td>0.3</td>
</tr>
<tr>
<td>Anderson et al. (1960) (1953 data)</td>
<td>0.4</td>
</tr>
<tr>
<td>Anderson et al. (1960) (1958 data)</td>
<td>0.2</td>
</tr>
<tr>
<td>Other</td>
<td></td>
</tr>
<tr>
<td>USPHS (1960) (physician visits)</td>
<td>0.1</td>
</tr>
<tr>
<td>USPHS (1960) (dental visits)</td>
<td>0.8</td>
</tr>
<tr>
<td>AMA (1978) (dental expenses)</td>
<td>1.0–1.7</td>
</tr>
<tr>
<td>Andersen and Benham (1970) (physician expenses)</td>
<td>0.4</td>
</tr>
<tr>
<td>Andersen and Benham (1970) (dental expenses)</td>
<td>1.2</td>
</tr>
<tr>
<td>Silver (1970) (physician expenses)</td>
<td>0.85</td>
</tr>
<tr>
<td>Silver (1970) (dental expenses)</td>
<td>2.4–3.2</td>
</tr>
<tr>
<td>Newman and Anderson, 1972 (dental expenses)</td>
<td>0.8</td>
</tr>
<tr>
<td>Feldstein (1973) (dental expenses)</td>
<td>1.2</td>
</tr>
<tr>
<td>Scanlon (1980) (Nursing Home expenses)</td>
<td>2.2</td>
</tr>
<tr>
<td>Sunshine and Dicker (1987) (dental expenses)</td>
<td>0.7–1.5</td>
</tr>
<tr>
<td>Hahn and Lefkowitz (1992) (dental expenses)</td>
<td>1.0</td>
</tr>
<tr>
<td>AHCPR (1997) (dental expenses)</td>
<td>1.1</td>
</tr>
<tr>
<td>Parker and Wong (1997) (Mexico, total expenses)</td>
<td>0.9–1.6</td>
</tr>
<tr>
<td>REGIONS [intermediate]</td>
<td></td>
</tr>
<tr>
<td>Feldstein (1971) (47 states 1958–1967, Hospital)</td>
<td>0.5</td>
</tr>
<tr>
<td>Fuchs and Kramer (1972) (33 states 1966, Physician)</td>
<td>0.9</td>
</tr>
<tr>
<td>Levit (1982) (50 states 1966, 1978, Total)</td>
<td>0.9</td>
</tr>
<tr>
<td>McLaughlin (1987) (25 SMSAs 1972–1982 Hospital)</td>
<td>0.7</td>
</tr>
<tr>
<td>Baker (1997) (3073 US counties 1986–1990, Medicare)</td>
<td>0.8</td>
</tr>
<tr>
<td>Di Matteo and Di Matteo (1998) (10 Canadian provinces 1965–1991)</td>
<td>0.8</td>
</tr>
<tr>
<td>NATIONS [macro]</td>
<td></td>
</tr>
<tr>
<td>Abel-Smith (1967) (33 countries, 1961)</td>
<td>1.3</td>
</tr>
<tr>
<td>Kleinman (1974) (16 countries, 1968)</td>
<td>1.2</td>
</tr>
<tr>
<td>Newhouse (1977) (13 countries, 1972)</td>
<td>1.3</td>
</tr>
<tr>
<td>Maxwell (1981) (10 countries, 1975)</td>
<td>1.4</td>
</tr>
</tbody>
</table>
analyses) is positively biased due to the use of observations that do not have closed financial boundaries. Hence, consumption elasticities estimated with observations across individuals or groups are typically mis-specified, and yield coefficients below 1.0 even when the true elasticity is 1.0 or above.

A somewhat similar issue arises on the supply side due to joint production costs. While easy to measure in total, costs for individual patients are not well-defined. This is most evident in pharmaceuticals, where the bulk of the costs arise in research and development, and “price per pill” has almost no relation to marginal cost. Jointness is also present in capital construction, emergency rooms, infection control, specialty training, quality assurance and indeed to some degree in most medical services. Decompose total expenditures \( X \) as \( X_{\text{group}} = \sum x_{i(\text{direct})} + X_{\text{joint}} \). To the extent that expenditures are for joint products, then in effect, purchasing is collectivized for the group (and thus should reflect group mean income) regardless of whether individual payments are made as insurance premiums, or based on per unit charges, surcharges on other medical costs, general tax contributions, percentage of wages, or other arbitrary allocation.

5. Concluding remarks

Units of observation must correspond to the units at which decisions are made. To the extent that costs are a result of expenditure decisions made at multiple levels, then a multilevel analysis is required. The consequences of income for spending are established at the level where budget constraints are fixed. For some types of health care, this is the individual household, for others, the hospital, region or insurance plan, but for much of medicine, it is the national budget constraint that is most relevant. With insurance, it is the average income of the group, and the fraction that the group is collectively willing to spend on medical care, that determines the health care budget, not the income of the particular patient being treated. Since the income elasticity varies with the level of analysis, health care is neither “a necessity” nor “a luxury” but “both”.

It is the boundaries of transactions, not geography, that are most relevant for the analysis of costs. Recognition of budget constraints leads to a realization that the effects of many individual and organizational variables will be allocative,
distributing the total across patients by need and other characteristics, rather than additive, summing upwards to some flexible aggregate. The observation that the amount spent on health care is determined by the amount available to spend rather than the amount of disease is not particularly new, and may even border on the obvious, but has not always been consistently incorporated in discussions of health policy, or in structuring the models used for empirical analysis of health spending. To the extent that group decisions become more important determinants of average expenditures than independent individual decisions, multilevel modeling becomes both efficient and necessary. In order to be empirically consistent and conceptually complete, midlevel analyses of regulations, organizations and technologies must link to a framework that models both micro and macro levels, and the dynamics of the connection between them.

Acknowledgements

This paper benefited from suggestions made by Sheila Smith, Mark Freeland, David B. Smith and Jacqueline Zinn and comments in workshops at the HCFA Office of the Actuary, the Johns Hopkins University Department of International Health, and the International Symposium on National Health Accounts in Rotterdam, June 1999.

References


Falk, I.S., Klem, M.C., Sinai, N., 1933. The Incidence of Illness and the Receipt and Costs of Medical Care Among Representative Family Groups. Committee on the Costs of Medical Care Publication #26, University of Chicago Press, Chicago.


