Forecasting Expenditures on Health Care in Developing Countries: An Econometric Approach

By

A.K. Nandakumar, PhD
Brandeis University

Jonathan Wilwerding, PhD
Abt Associates, Inc.

This work on this paper was funded by the Partners for Health Reformplus (PHRplus) Project funded by USAID under contract # HRN-C-00-00-00019-00 and implemented by Abt Associates Inc. and partners Development Associates Inc,; Emory University Rollins School of Public Health; Program for Appropiate Technologies in Health; Social Sectors Development Strategies, Inc; Training Resources Group; Tulane University School of Public Health and Tropical Medicine; and University Research Co.; LLC

The views expressed in this paper are those of the authors and any errors and omissions are entirely theirs

This technical note has been prepared for the IAAHS Colloquim in Dresden Germany, April 28-29, 2004.

Forecasting Expenditures on Health Care in Developing Countries: An Econometric Approach

I. Introduction

In the next fifty years, the share of the world population aged 65 or older will double, the average age will rise from 26.5 to 36.2, and the fast-growing 80+ age group's share will quadruple rising from 3% to 8% (UN Population Division 1998). What is more the process is speeding up. Today, about two-thirds of all older people are living in the developing world; by 2025, it will be 75%. Although some countries could be considered to still be in the pre-transitional phase of the demographic transition with high fertility and mortality rates, in most developing countries there will be a very significant increase in the size of the population over 65 in the next 25 years. Medical advances, improved diet and decreased fertility are the root causes of the aging boom the world's elderly in developing countries by 2050. This will put severe pressure on public finances and policy responses will have to be both radical and relatively swift[1].

Virtually all of the available literature on the costs of aging pertains to the developed countries, which form part of the Organization of Economic Co-operation and Development (OECD), and the bulk comes from the USA. These studies tend to fall into two categories: the actuarial approaches and the micro-simulation models[2].

At this stage, we simply do not know what proportion of total health expenditures in developing countries is spent on the aged or if that proportion is rising or falling. There are good reasons to think that it will rise, and rise sharply because of the very rapid increase in the number of old people. It is critical to be able to predict what the impact of these changes is likely to be on total health expenditures, and on the demands for public sector care.

Estimating health expenditures in developing countries poses a unique set of methodological challenges. To start with the kind of longitudinal data sets available in developed countries are not available. Further, increases in health expenditures on the elderly will be affected by both the demographic and epidemiological transitions. Very few studies exist on disability levels (a good predictor for long-term care) among the

[1] Bloom, D.E., Nandakumar, A.K., Bhawalkar, M., "The Demography of Aging in Japan and the US," Chapter in Book based on International Symposium on "Aging and Health: Environment, Work, and Behavior" September 2001 at the American Academy of Arts and Sciences in Cambridge. The book is in print with Harvard Press

[2] Mahal, A., Berman, P., "Health Expenditures and the Elderly: A Survey of Issues in Forecasting, Methods Used, and Relevance for Developing Countries," Global Burden of Disease in Aging Populations, Research Paper No. 01.23

elderly though some studies suggest that the elderly in developing countries report higher rates of Activities of Daily Living limitations than their counterparts in the US[3].

Health systems in developing countries also differ in many fundamental ways from those of developed countries. National Health Accounts studies conducted in developing countries show financing in developing countries tends to be fragmented and that out-of-pocket expenditures are significantly larger in poor than in richer countries[4]. Moreover, it is possible that old people do not use formal acute care facilities to the same extent as younger people[5], and the extent of long term care in the formal health sector is negligible in most countries because of a reliance on family members.

II. The Methodology

The United States Agency for International Development (USAID) is interested in the impact of aging on funding for priority services like reproductive and maternal health, child health, HIV/AIDS, and infectious disease. This research is funded through USAID's PHR*plus* project and is focused on the development of a methodology to estimate the impact of aging populations on a developing country's health expenditures. PHR*plus* is working with University of Hawaii's East West Center (EWC) and counterparts in Jordan and the Philippines on this analysis. In this technical note we present an econometric approach to modeling health expenditures on the elderly in developing countries.

We decided to use a National Health Accounts (NHA) framework for the classification of expenditures. There were a number of reasons for this decision including:

1. The NHA framework provides estimates for total spending on health care by sources, financing intermediaries, and end uses.
2. The framework is rigorous and provides for validation of data
3. It permits the analysis of expenditures on the elderly to be done in the context of the larger health system.

The Guide to Producing National Health Accounts released by the World Bank, World Health Organization and USAID in 2003 describes NHA as "A systematic, comprehensive, and consistent monitoring of resource flows in a country's health system."

---

[3] Nandakumar, A.K., El-Adawy, M., Cohen, M., "Perception of Health Status and Limitations in Activities of Daily Living among the Egyptian Elderly," Data for Decision Making Project, December 1998. Funded by the United States Agency for International Development

[4] Nandakumar, A.K., Bhawalkar, M., Tien, M., Ramos, R., Susna, D., "Synthesis of Findings from NHA Studies in Twenty Six Countries," PHRplus Project, Bethesda, MD, Abt Associates, February 2004 (Draft)

[5] Nandakumar, A.K., Wilwerding, J., Bhawalkar, M., Arbaji, A., "Jordan Healthcare Utilization and Expenditure Survey, 2000," Primary Health Care Initiative Project, Bethesda, MD, Abt Associates, 2002

The National Health Accounts estimates are supplemented with budgetary data, facility level costing studies, and National Household Health Care Utilization and Expenditure Surveys.

We estimate expenditures on the elderly in four steps:

1. Estimate expenditures for the base year, using household survey data and costing studies
2. Use a model of macro-economic growth to obtain upper bound estimate on public expenditures in projection year
3. Age survey data from the base year to obtain data for the projection year
4. Apply econometric models to the aged data to obtain estimated expenditures for the projection year

In the following sections we describe these four steps in greater detail

III. Estimates of Total Expenditures and Use of Healthcare in the Base Year

Total expenditures on healthcare in a given year are the sum of total public expenditures, total private expenditures and total expenditures by donors. We take these up *seriatim*.

A. Total Public Expenditures

To estimate total expenditures on healthcare, the first step is to say how we estimate total public expenditures. We begin with some notation.

(1)

$$y_{ijkl}$$

$$q_{ijkl}$$

$$c_{jkl}$$

$$i = \{1, 2, ..., I\}$$
$$j = \{1, 2, ..., J\}$$
$$k = \{1, 2, ..., K\}$$
$$l = \{1, 2, ..., L\}$$

We define y_{ijkl} as total spending by entity *l* for care of type *j* for individual *i* in year *k*. The subscript *i* indexes individuals while *j* indexes *functions.* Functions are types of service, such as outpatient and inpatient care. The subscript *k* indexes years. Because we are concerned with the base year at the moment, *k* would have a values corresponding to that year. If the year 2000 were the base year, for example, y_{ijkl} would be total spending

by *l* for care of type *j* for individual *i* in the year 2000. Finally, the subscript *l* indexes types of public entity. Where *l* is a Ministry of Health, for example, and *j* outpatient care, y_{ijkl} is the amount paid by the Ministry for outpatient care of *i* in year *k*.

The variable q_{ijkl} is the quantity of care of type *j* used by individual *i* in year *k* provided by entity *l*. For outpatient care obtained at facilities funded by a Ministry of Health facilities, for example, q_{ijkl} is the number of outpatient visits by *i* in year *k* to a Ministry facility.

Finally, c_{jkl} is the unit cost of care of type *j* –e.g. outpatient or inpatient—to entity *l* for care provided by *l* in year *k*. The unit cost of care may vary across public entities. The cost per visit of outpatient care to a Ministry of Health facility, for example, may differ from the cost per visit for outpatient care paid by other public entities.

Note that a public entity pays only part of the cost of care at its facilities because patients frequently make out of pocket payments for aspects of their care, such as pharmaceuticals. The latter is the private, household or out of pocket spending that we considered elsewhere.

B. Public Expenditures on Healthcare in the Base Year, Total and Disaggregated by Public Entity

Let *l* be a Ministry of Health, for example. The total cost paid by a Ministry of Health for any individual in a given year is the amount of Ministry care used by that individual, multiplied by the cost per visit, or per admission. For outpatient care, this is the product:

(2) $$\left(q_{i,outpatient,k,l} \times c_{outpatient,k,l}\right)$$

For inpatient care, the counterpart expression is

(3) $$\left(q_{i,inpatient,k,l} \times c_{inpatient,k,l}\right).$$

Let Y_{kl}^{g} be total public expenditures by entity *l* in year *k*. We obtain this by summing (2) and (3) across types of care—e.g. inpatient and outpatient—as well as individuals.

(4) $$Y_{kl}^{g} = \sum_{j=1}^{J} \sum_{i=1}^{I} \left(q_{ijkl} \times c_{jkl}\right)$$

Total public expenditures in a given year, then, are the sum across spending by all individuals, functions and public entities:

(5) $$Y_{k}^{g} = \sum_{l=}^{L} \left(\sum_{j=1}^{J} \sum_{i=1}^{I} \left(q_{ijkl} \times c_{jkl}\right)\right)$$

Let Q_{jkl} be the sum of q_{ijkl} across individuals. We can then rewrite (5) as

(6) $$Y_k^g = \sum_{l=1}^{L} \left(\sum_{j=1}^{J} \left(Q_{jkl} \times c_{jkl} \right) \right)$$

To estimate equation total public expenditures by means of equation (6) for the base year, we need estimates of Q_{jkl} and c_{jkl} for that year—i.e. for *k = base year*. We obtain c_{jkl} from costing studies. We esimate Q_{jkl} using household survey data and econometric methods.

C. Estimating Q_{jkl}

We estimate Q_{jkl} in two steps, both of which use household survey data. In the fist step, we estimate total outpatient visits and total inpatient admissions. These totals implicitly sum across visits or admissions to facilities of all types. For this reason, we suppress the subscript *l* in the expressions that follow.

In the second step, we allocate total visits and admission across provider types. Here we use household survey data as well, but in this case, for each public entity, we estimate the likelihood that an individual will seek care from that entity. We do this for both outpatient visits and inpatient admissions, estimating a separate model for each case. To estimate these probabilities, we estimate multinomial logit models (MNLM) of choice of provider.

D. Step 1 in Estimating Q_{jkl} *: Total Outpatient Visits in the Base Year*

To estimate total utilization of outpatient care—i.e. outpatient visits—in the base year, we begin with outpatient utilization by individual *i* in year *k:*

(6) $$q_{i,outpatient,k}$$

We obtain total outpatient utilization for the sample in year *k* by summing outpatient visits across individuals.

(7) $$q_{outpatient,k} = \sum_{i=1}^{N} w_i q_{i,outpatient,k}$$

This allows us to estimate average outpatient use in year *k* as above, by dividing total outpatient visits by the sample size.

(8) $$\overline{q_{outpatient,k}} = 1/N \sum_{i=1}^{N} w_i q_{i,outpatient,k}$$

Finally, total outpatient visits for the population in year *k* we obtain as the product of average visits multiplied by the population.

(9) $$\overline{Q_{outpatient,k}} = \left(1/N \sum_{i=1}^{N} w_i q_{i,outpatient,k} \right) \times population_k$$

Note that this step is necessary only where sample weights are normalized. Household survey weights sometimes are normalized. Where the w_i are non-normalized probability weights, equation (8) will suffice.

D. Step 1 in Estimating Q_{jkl}: Total Inpatient Admissions in the Base Year

We estimate utilization of inpatient care in an exactly analogous way. Utilization by individual *i* in year *k* is

(10) $$q_{i,inpatient,k}$$

We estimate inpatient admissions across individuals to obtain total inpatient use in year *k* for the sample.

(11) $$q_{inpatient,k} = \sum_{i=1}^{N} w_i q_{i,inpatient,k}$$

Dividing by the sample size (8,306), we obtain average inpatient use for the sample in year *k*.

(12) $$\overline{q_{inpatient,k}} = 1/N \sum_{i=1}^{N} w_i q_{i,inpatient,k}$$

Finally, total inpatient use for the population of Jordan in year *k* we estimate as average use multiplied by the population of Jordan in year *k*.

(13) $$\overline{Q_{inpatient,k}} = \left(1/N \sum_{i=1}^{N} w_i q_{i,inpatient,k} \right) \times population_k$$

Note, once again, that this step is necessary only where sample weights are normalized. Where the w_i are non-normalized probability weights, equation (12) will suffice.

E. Step 2 in Estimating Q_{jkl}: Allocating Visits and Admissions to Public Entities

Thus far, we have estimates of Q_{jk} by means of expressions (9) and (13) above—these give Q_{jk} for *j = inpatient* and *j = outpatient*. To estimate Q_{jkl}, we need to allocate total inpatient and outpatient visits across public entiites. Let p_{jkl} be the probability that a function of type *j*—i.e. an outpatient visit or inpatient admission—will occur at a public facility of type *l*—e.g. a facility operated by a Ministry of Health. Then

(14) $$Q_{jkl} = p_{jkl} \times Q_{jk}$$

is an esitmate of toal visits or admissions to facilities of type *l* in year *k*, in this case the base year. To allocate visits and admissions across provider types, then, it remains to estimate p_{jkl} for each public entity, *l*.

We assume a multinomial logit model of p_{jkl}. Let *p* indicate choice of provider. This variable takes values $p = \{1, 2, ..., L\}$ just as the index *l* does above. The MNLM assumes that:

(15)

$$P(p_{ijk} = m \mid x_i) = \frac{e^{\mathbf{x}_i'\boldsymbol{\beta}_m}}{1 + \sum_{l=2}^{L} e^{\mathbf{x}_i'\boldsymbol{\beta}_l}}$$

Note that, in effect, we estimate this equation twice, once for outpatient visits and once for inpatient admissions.

The vector **x** includes predictors of choice of provider. In our models, these include, for example, individuals' insurance coverage, per capita household income and urban (vs. rural) location. Insurance coverage, or the lack of it, may reduce the price of care at one type of provider relative to others, or may bar use of some care completely. Income affects choice of provider to the extent that the price of care varies by provider type. Location affects choice of provider as well, in cases where some providers are located predominantly in some areas but not others. Apart from these variables, we include standard demographic variables, including age, gender, marital status, employment and education.

Once we eatiamte the p_{jkl} and Q_{jk} for the base year, we obtain Q_{jkl} as the product

(16) $$Q_{jkl} = p_{jkl} \times Q_{jk}$$

for each value of *l;* this give total public expenditures for each type of provider. By applying equation (6), we complete our estimate of total public expenditures by public entity.

Note in this connection that household survey data typically underestimates Q_{jkl} for the case of inpatient care. Many household surveys use a twelve-month recall period for questions regarding inpatient care. Subjects' reports are not reliable for periods of this length, and it is necessary to adjust estimates of Q_{jkl} using secondary data sources, including official admission statistics, for example.

E. Private Expenditures in the Base Year

Total expenditures on health care in the base year are the sum of total public expenditures, total private expenditures and total expenditures by donors. Thus, to estimate total expenditures overall, we need an estimate of total private expenditures.

Abusing our notation slightly, we redefine y_{ijk} as household expenditures on individual *i* of type *j* in year *k*. As before, the subscript *i* indicates individual, *j* indicates functions, and *k* the year in which *i* used function *j*. Expenditures—i.e. functions—may be any of four types: outpatient care, inpatient care, transportation to outpatient care and expenditures on pharmaceuticals.

Total private, out of pocket expenditures on health care for an individual in a given year are the sum of that individual's expenditures across function—e.g. outpatient care, inpatient care, transportation to outpatient care and routine pharmaceuticals. This is equivalent to summing y_{ijk} across types of expenditures *j*.

(2) $$y_{ik} = \sum_{j=1}^{J} y_{ijk}$$

To obtain total expenditures for the sample for a given year *k* we sum total expenditures across individuals in year *k*.

(3) $$y_k = \sum_{i=1}^{N} w_i y_{ik}$$

We obtain an estimate of average total expenditures by individuals in the sample for year *k* by dividing total expenditures by the sample size, *N*.

(4) $$\overline{y_k} = 1/N \sum_{i=1}^{N} w_i y_{ik}$$

Finally we estimate total expenditures for the population for a given year, by multiplying average individual expenditures by the size of the population in that year.

(5) $$Y_k^{private} = \overline{y_k} \times populaiton_k$$

Note that step (4) above will suffice where weights are not normalized.

F. Expenditures by Donors in the Base Year

Contrary to developed countries, many low and middle income countries are highly dependent on external assistance. National Health Accounts studies have shown that in some instances external assistance can be as high as a third of total health expenditures of a country. External assistance as a percentage of total health expendiutres is inversely correlated to a country's per capita income[6].

The only way in which to get accurate information on external assistance is to obtain this information from both government records and through a survey of donors. Where longitudinal data on donor assistance to a low and middle income country is available one can model it as a function of economic growth. In the case of middle income countries external assistance tends to be less than five percent of total health expenditures.

IV. Modeling macro-economic growth

We use a sophisticated, recent model to estimate macroeconomic growth[7]. This model assumes that every country has a ceiling on the level of per capita income that it can attain, which is determined by such characteristics as its geography, natural resources, public policies, and human capital. Every country also is assumed to be out of equilibrium with respect to its attainable level of income, but is always tending toward that level. Within this framework, if two countries with the same starting level of income were compared, the country with the higher potential income would grow more quickly. If all countries had the same underlying characteristics the per capita income in each country would converge toward the same level.[8]

This empirical framework is flexible and permits testing of alternative hypotheses about the determinants of economic growth. For example, a series of empirical studies have shown that initial life expectancy has a positive, sizable, statistically significant, and independent influence on the pace of subsequent economic growth. Indeed, about a

[6] Nandakumar, A.K., Bhawalkar, M., Tien, M., Ramos, R., Susna, D., "Synthesis of Findings from NHA Studies in Twenty Six Countries," PHRplus Project, Bethesda, MD, Abt Associates, February 2004 (Draft).

[7] Bloom, D.E., Canning, D., Nandakumar, A.K., Sevilla, J., Huzarski, K., Levy, D., Bhawalkar, M., "Demographic Transition and Economic Opportunity: The Case of Jordan," PHR*plus* Project, Bethesda, MD, Abt Associates, 2001.

[8] See Robert J. Barro, "Economic Growth in a Cross Section of Countries", *Quarterly Journal of Economics,* 106 (1991): pp. 407-443; Robert J. Barro and Xavier Sala-i-Martin, *Economic Growth* (New York: McGraw Hill, 1995).

quarter of Africa's slow rate of economic growth (relative to East Asia) may be due to the low level of life expectancy there.[9]

More recently, researchers introduced the rate of population growth and changes in the dependency burden into this class of empirical growth models.[10] In each country population growth has been shown to affect economic growth. It is not so much the overall rate of population growth but rather changes in the dependency ratio that matters when it comes to economic growth.

Given different assumptions of macro-economic growth we also model overall government expenditures. In the next step given overall government expenditures we model the share of expenditures going to health care by type of public entity. These estimates provide an upper bound against which the estimates of expenditures using individual level information are compared.

V. Estimates of Expenditures in the Forecast Year

A. Overview

To estimate expenditures for the forecast year, we proceed in four steps. First, we estimate econometric models of private expenditures, visits, admissions, and choice of provider using data for the base year. Second, we "age" the data from the base year to represent the forecast year. For each individual and variable in our econometric models, we predict values that the same individual would have in the forecast year. Third, we apply the estimated econometric models to the aged data set, to predict private expenditures, visits, admissions, and choice of provider for the forecast year. This gives us private expenditures as well as total visits and admissions by provider type. Finally, we use a macroeconomic model to inflate the unit cost of care at each public entity. This allows us, in turn, to estimate total public expenditures.

B. Econometric Models Estimated for the Base Year

In the first step, we estimate models of private out of pocket expenditures, utilization and choice of provider using data household survey data from the base year. We intend to apply these models to data for the forecast year as well, so we present them in their general form, where the index *k* can take either the base or forecast years.

Private expenditures

Let Y_k^p be private out of pocket expenditures in year *k.* We estimate expected private expenditures in the year *k* as the product.

[9] Bloom and Sachs, 1998.
[10] Bloom and Williamson, 1997 and 1998; Bloom and Sachs, 1998; Bloom and Malaney, 1998; and Bloom, Canning, and Malaney, 2000.

(19) $$E\left(Y_k^p\right) = P\left(Y_k^p > 0\right) \cdot E\left(Y_k^p \mid Y_k^p > 0\right)$$

In expression (19) $P\left(Y_k^p > 0\right)$ is the probability that expenditures are greater than zero in the base year. We estimate $P\left(Y_k^p > 0\right)$ by logistic regression:

(20) $$P\left(Y_k^p > 0 \mid X_k\right) = \frac{e^{\mathrm{X_k}\beta}}{1 + e^{\mathrm{X_k}\beta}}$$

To obtain the product in expression (19), we estimate $E\left(Y_k^p \mid Y_k^p > 0\right)$ by OLS:

(21) $$E\left(Y_k^p \mid Y_k^p > 0\right) = \mathrm{x_k}\beta + \varepsilon$$

We attempted other modeling strategies for private expenditures. Initially, it seemed as though the natural model for expenditures on healthcare would satisfy assumptions of the Tobit model. The latter model, however, failed to predict well within sample. Selection (e.g. Heckman) models also predicted poorly within sample.

In equation (20) we require a vector of variables that predict the likelihood that an individual would have greater than zero out of pocket expenditures on healthcare. We use three principal such predictors in the baseline data. First, an individual will be more likely to incur out of pocket expenses if she sought outpatient care. Household surveys typically ask individuals whether they visited any health service provider such as a doctor, pharmacist nurse or pharmacist in order to treat a health condition. Second, one would expect an individual to be less likely to incur out of pocket expenses if she were covered by insurance. Household surveys also typically include indicators of insurance coverage. Finally, in preliminary estimates using actually household survey data from the nation of Jordan, we found that employment is associated with the probability of having non-zero out of pocket expenditures.

We estimated equation (20) using specifications where the vector of explanatory variables also includes variables for age, gender, marital status, education, income, urban residence, self-reported health status, a diagnosis of serious or chronic illness. In preliminary estimates, these variables were not statistically associated with the probability of non-zero health expenses (Likelihood Ratio test: Prob. > F = 0.85).

With Equation (21) we seek to estimate expected out of pocket expenditures, conditional on a second vector of predictors, for those with non-zero expenses. Expenditures are naturally associated with the i) the overall cost of care and ii) individuals' ability and willingness to pay for care. As proxy measures for the cost of care, we use age, gender and severity of illness. These are proxy indicators of how much care an individual is likely to need. We also include an indicator of urban location as a measure of availability and, perhaps, cost of care. As measures of resources and willingness to pay, we include measures of income and insurance coverage. Note that there are likely to be income effects on demand for healthcare.

Outpatient visits (curative)

Visits to outpatient facilities are either curative or preventive. Household survey data does not always include data on preventive care. It does, however, typically allow us to estimate outpatient visits for curative care, including visits for antenatal medical problems. Hereafter, we use 'outpatient visits' to refer to curative outpatient visits.

Many surveys ask subjects, first, to report whether they had an illness during the reference period of the survey. If they reported no illness, they were asked no further questions about outpatient care. We assume that all such persons had zero curative visits to an outpatient facility. We define a variable indicating whether an individual had an illness in a give year:

(22) $$S_{i,k} = 1\left(i\ was\ ill\ in\ year\ k\right)$$

For those who report having an illness, household surveys typically ask subjects to report the number of visits that they made to outpatient facilities during a specified reference period of the study. Let the reference period be *m* weeks; we assume that visits have been annualized by mutiplying this number by 52/*m*.

We define a variable for the number of visits or admissions by an individual *i* in a given year.

(23) q_{ijk} = Visits of type *j* by individual *i* in the year *k*

Where *j* has the value 'outpatient' and *k* is the base year, this variable counts outpatient visits by individual *i* in the latter year.

We obtain outpatient visits in a given year, as the product of i) the probability that an individual had an illness in that year and ii) the expected number of outpatient visits, for people that had an illness.

(24) $$E\left(Q_{jk}\right) = P\left(S_k = 1\right) \cdot E\left(Q_{jk} \mid S_{,k} = 1\right)$$

We estimate $P\left(S_k = 1\right)$ using a logistic regression model, as in (20).

Where the index *j* is set to 'outpatient', $E\left(Q_{,jk} \mid S_{,k} = 1\right)$ is the expectation of outpatient visits in year *k*, conditional on having an illness in that year. Note, first, that we are less interested in the specific distribution of outpatient visits than we are in identifying general properties of this distribution—e.g. its mean. The model that we estimate, therefore, is a reduced form models of outpatient visits (and, later, inpatient admissions).

An individual can have only a non-negative integer number of outpatient visits in a given year. For this reason, we begin by assuming that the data are generated by a Poisson process.

In fact, there may be two processes at work. A first process separates individuals who always have zero counts even if they are ill. In our case, these may be individuals who do not seek medical care regardless of ill health—e.g. perhaps because they are either too poor or too distant from medical care. A second process determines the number of outpatient visits amoung those who do not always have zero counts.

In light of these two processes, we estimate outpatient visits in a given year, conditional on having an illness, by means of a zero-inflated negative binomial regression model (ZINB). Estimation proceed by maximizing the log-likelihood L defined in expression (25).[11][12]

(25)

$$m = 1/\alpha$$

$$p_i = 1/\left(1+\alpha\mu_i\right)$$

$$\xi_i^{\beta} = \mathbf{x}_i'\boldsymbol{\beta} + offset_i^{\beta}$$

$$\xi_i^{\gamma} = \mathbf{z}_i'\boldsymbol{\gamma} + offset_i^{\gamma}$$

$$\mu_i = \exp\left(\xi_i^{\beta}\right)$$

$$L = \sum_{i \in S} w_i \ln\left[F\left(\xi_i^{\gamma}\right) + \left\{1 - F\left(\xi_i^{\gamma}\right)\right\} p_i^{m}\right]$$

$$+ \sum_{i \notin S}\left[w_i \ln\left\{1 - F\left(\xi_i^{\gamma}\right)\right\} + w_i \ln\Gamma\left(m + y_i\right) - w_i \ln\Gamma\left(y_i + 1\right) -\right.$$

$$\left. w_i \ln\Gamma\left(m\right) + w_i\, m \ln p_i + w_i\, y_i \ln\left(1 - p\right)\right]$$

[11] Here we are interested in modeling outpatient visits for the year 2000. For this reason, we suppress indexes *j* and *k* to simplify our exposition.
[12] StataCorp. 2003. *STATA Statistical Software: Release 8.0, Reference Manual,* v.4 p.338. (College Station, TX: Stata Corporation).

The vector ***Z*** includes predictors of the probability that an individual will always have a zero count. In our case, these are predictors that a person will resist using healthcare even if ill. Natural barriers to using healthcare are income, distance from healthcare, and perhaps attitudes that make individuals distrustful of healthcare providers. Household surveys generally collect data regarding income and urban (as against rural) location. These may proxy for income and distance, respectively. As proxy for attitudes regarding healthcare, we include a measure of individuals' age and completed level of education.

The vector ***X*** includes predictors of outpatient visits among those who may not have a zero count. One would expect annual outpatient visits to be associated, first, with health status. More seriously or more chronically ill individuals should use more outpatient care. Household survey data typically include a number of proxy measures for health status including age and self-reported health status relative to age peers. They also include variable indicating whether an individual has ever been diagnosed with a serious illness, has a chronic condition, or has an illness sufficient to restrict activities of daily living.

One would expect use of outpatient care also to be associated with resources. For this reason, we include among the vector ***X*** variables measuring income, urban residence and possession of insurance. We also include a measure of individuals' level of education, age, gender and marital status.

In preliminary estimates, we find that that the ZINB model fits the data better than either any Poisson regression model, and better than an unmodified negative binomial regression model (by Vuong test).

Choice of provider

We estimate total outpatient visits, in order to estimate total public expenditures. To estimate public expenditures, we need i) to allocate outpatient visits across providers, ii) to estimate the public cost per visit for each provider and iii) to then make use of equation (6) above. After estimating total outpatient visits, we allocate them across providers by estimating the probability that someone who seeks care will choose one provider's facilities over another. Once again, we model these probabilities using a multinomial logit model (MNLM), as in the case of the base case.[13]

Let p be a variable indicating choice of provider. This variable takes values corresponding to public and private providers—these are values of the index l that we use above. We model probability that p takes a given value m as follows.

[13] See, for example McFadden (1973). For a recent exposition, see Long (1997).

(26)

$$P(p_i = m | x_i) = \frac{e^{\mathbf{x}_i'\boldsymbol{\beta}_m}}{1 + \sum_{l=2}^{L} e^{\mathbf{x}_i'\boldsymbol{\beta}_l}}$$

The vector **x** here includes several predictors of choice of provider. These are essentially the same as the set of variables that we indicate in connection with expression for the base case.

Inpatient admissions

We estimate public expenditures on inpatient care in a way analogous to our method for estimating public expenditures on outpatient care. First, we estimate total inpatient admissions, in a manner analogous to equation (24). Next, we allocate admissions across public provider, using probabilities estimated in a way analogous to equation (26). Finally, we multiply total admissions at public entities, by their respective costs, as estimated for the year 2000 using cost studies and inflated for the year 2015 using GDP inflators derived from macro-economic projections.

D. Method for Aging Bas Year Data to Represent the Forecast Year

In the second step of our forecasting method, we "age" survey data from the base year to obtain "survey data" for the forecast year. It is sufficient here to age only data for variables that appear in econometric models that we use to predict use and expenditures.[14] We outline our method here.

Age

For each observation household surveys typically include a recorded age. For all observations in the data set we begin by assigning an age equal to age in the base year plus fifteen—we are assuming a fifteen-year forecast. While this transformation is straightforward, note that it yields a data set in which no one is under the age of fifteen. To solve this problem, after aging the data we add observations to the data set, to reflect demographic projections regarding the size of the population under fifteen years of age in the forecast year.

To account for the effects of mortality in the aged data set, we weight the data using 15-year survival rates, by age, estimated using data from UN Population studies.

Marital status

There are two cases to consider in assigning marital status for the forecast year: i) individuals who were married or formerly married in the base year, and ii) individuals

[14] We assume that household survey data includes the variables that we discuss here.

who are single in the base year and created observations—these are observations with age less than fifteen in the forecast year.

We hold marital status constant for individuals that were either married or formerly married in the base year. For individuals that are single in the base year, and for created observations, we assign marital status randomly by age category, in a way that preserves the distribution of marital status that we observe in data from the base year. This procedure holds the distribution of marital status within each age group approximately constant between the base and forecast years.

Education

We proceed in a similar way for education and self-reported health status. For those over the age of 25 in the base year, we assume that their level of education does not change by the forecast year. For all other age groups, we begin by estimating the distribution of level of education within each age group. After aging the data we estimate the distribution of education. We then randomly assign individuals to levels of education by age group, where necessary, to reproduce the base year distribution.

Income and Employment

To predict income for the forecast year, we use the macroeconomic model described above to predict GDP for the forecast year, and U.N. data to predict population. The quotient of these is an estimate of per capita income for the forecast year. We have an estimate of average per capita income for the base year derived from household survey data. This allows us to derive a rate at which per capita income grows between the base and forecast years:

(27)

$$r_{base,\, forecast} = \frac{per\ capita\ income_{forecast}}{per\ capita\ income_{base}}$$

For each observation in the forecast year we predict per capita income as the product of per capita income for the base year and $r_{base,\, forecast}$.

We use econometric methods to predict employment for the forecast year. We treat full-time employment as a binary variable, $Y = \{0,1\}$, and model employment by logistic regression.

(28)

$$P\left(Y_{i,\, employed,\, k} = 1 \mid X_i\right) = \frac{e^{X_i' \beta}}{1 + e^{X_i' \beta}}$$

The vector of explanatory variables in expression (28) includes gender, age, the square of age, years of completed education, the square of years of completed education, and urban (vs. rural) location. We estimate this model using data for the base year, and then apply model parameters to predict employment status for the forecast year using the "aged" data set.

Health Status

We treat self-reported health status (relative to age peers) by random assignment alone. We begin by estimating the distribution of health status for the base year by age, gender and income quintile. We then reproduce this distribution for the forecast year by random assignment.

Insurance

We predict insurance coverage, overall and by type of insurance, by multinomial logistic regression. This requires several steps. First, in addition to predicting employment status by means of expression (28), we model marital status using data for the base year.

(29)

$$P\left(Y_{i,married,k}=1 \mid X_i\right)=\frac{e^{X_i'\beta}}{1+e^{X_i'\beta}}$$

The vector of explanatory variables in this expression, **x** includes gender, age, and the square of age.

We apply (28) and (29) to obtain predicted values for the base and forecast years.

(30)

$$\hat{P}\left(Y_{i,employed,k} \mid X_i\right)$$

$$\hat{P}\left(Y_{i,married,k} \mid X_i\right).$$

Next we estimate insurance coverage by multinomial logistic regression, using household survey data from the base year. The vector of explanatory variables in this regression includes the predicted probabilities from expression (30) where k is the base year. Finally, we predict insurance status for the forecast year using the predicted probabilities from expression (30) as explanatory variables, alongside age, the square of age, gender, education, income, health status and urban location.

C. Cost per Visit and Cost per Admission to Public Entities

We obtain the cost per visit for public entities from costing studies (see above). We inflate these costs using the growth rate projections obtained from the earlier macro-economic modeling. The total costs of outpatient visits and inpatient admissions at government facilities are compared against the upper bound for public budgets. As we recall these estimates were computed earlier.

VI. Application to Jordan

In Annex 1 we present a series of charts that present the results from the application of this methodology to Jordan. These estimates are still being refined and might change but are useful in presenting the direction of expected change. We modeled the projections under three scenarios of macro-economic growth. The macro-economic model of growth showed that, with proper policies in place, Jordan's economy had the potential to grow at 5.6% per annum between 2000 and 2015. We also modeled the scenario where the economy grew at an annual rate of 2% per year and 3% per year.

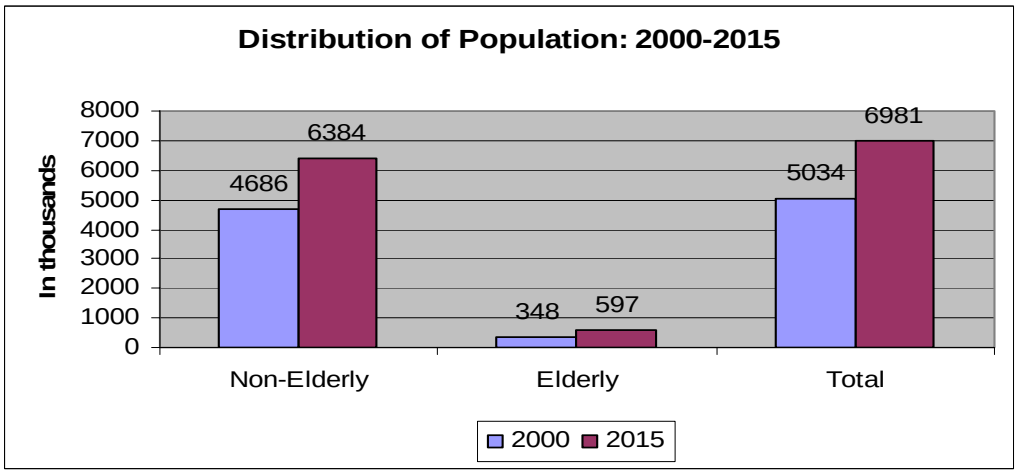
The elderly as a percentage of the population in Jordan are projected to increase from 7% in 2000 to 9% in 2015. Irrespective of the growth rate assumptions the use of public facilities will decline and use of private providers will increase both for outpatient visits and inpatient admissions. In 2000, the elderly accounted for 20.2% of total health expenditures in the country. This is projected to increase to 23.2% under a 5.6% per annum growth rate assumption, 32.7% under a 3% growth rate assumption, and 37.9% under a 2% growth rate assumption. Thus, in 2015 while the elderly will account for only 9% of the population they might account for over a third of national health expenditures. At the same time there will be a steep increase in the burden of out-of-pocket expenditures on the elderly especially for medicines. The analysis shows that the demand on the health system made by the elderly has the potential to adversely affect outlays for key public health services and maternal and child health services.

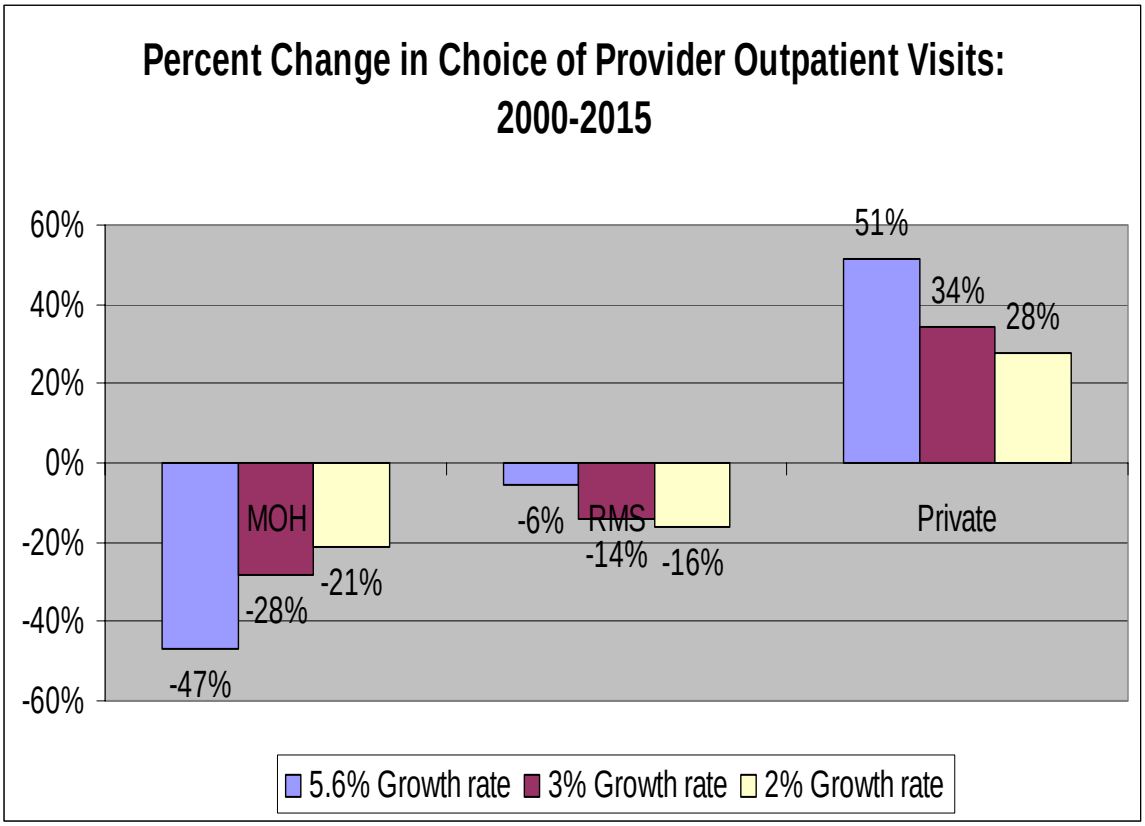
VII. Conclusion

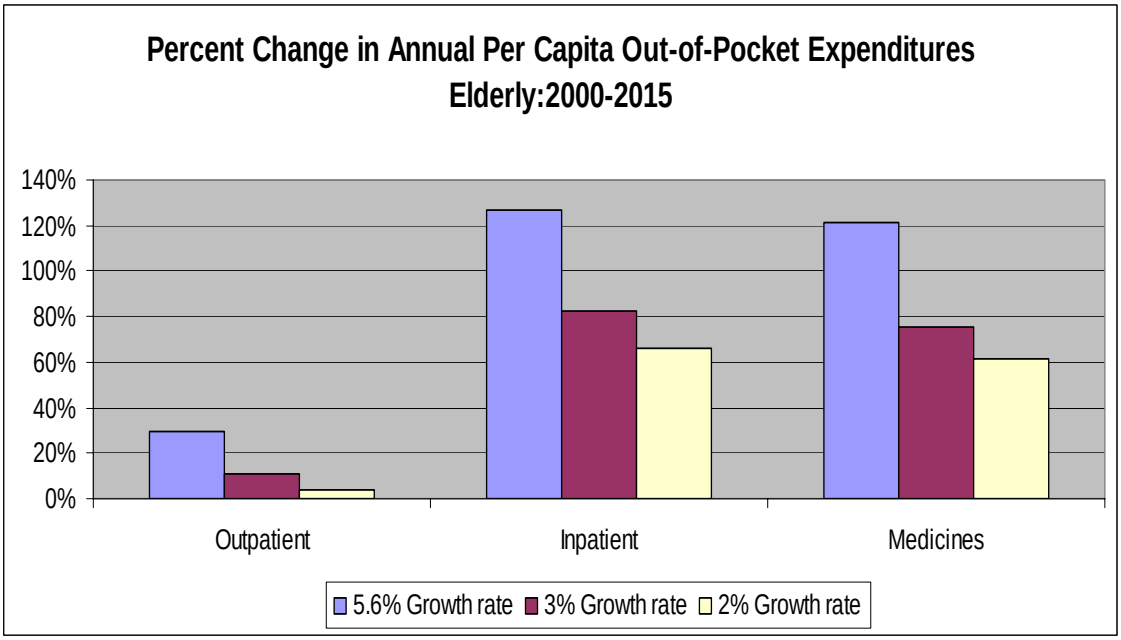
As countries age and develop the demand for health care services will grow rapidly, the demand for private sector services vis-à-vis public services will grow, competition for scarce health resources will increase, pressure on the non-health public sector will also increase, and the burden of out-of-pocket expenditures on the elderly will rise over time. Our analysis has shown that one needs to consider expenditures on the elderly within the context of the entire health system and not in isolation as has been done in other studies. This involves modeling all population sub-groups and not just the elderly. Next, we found that what happens in the public sector affects the private sector. Hence merely modeling private expenditures without taking into consideration public expenditures and changes can lead to erroneous conclusions. In the years to come governments in low and middle income countries will be forced to respond rapidly and swiftly to the pressures on the health system and the economy because of the increasing share of the elderly.

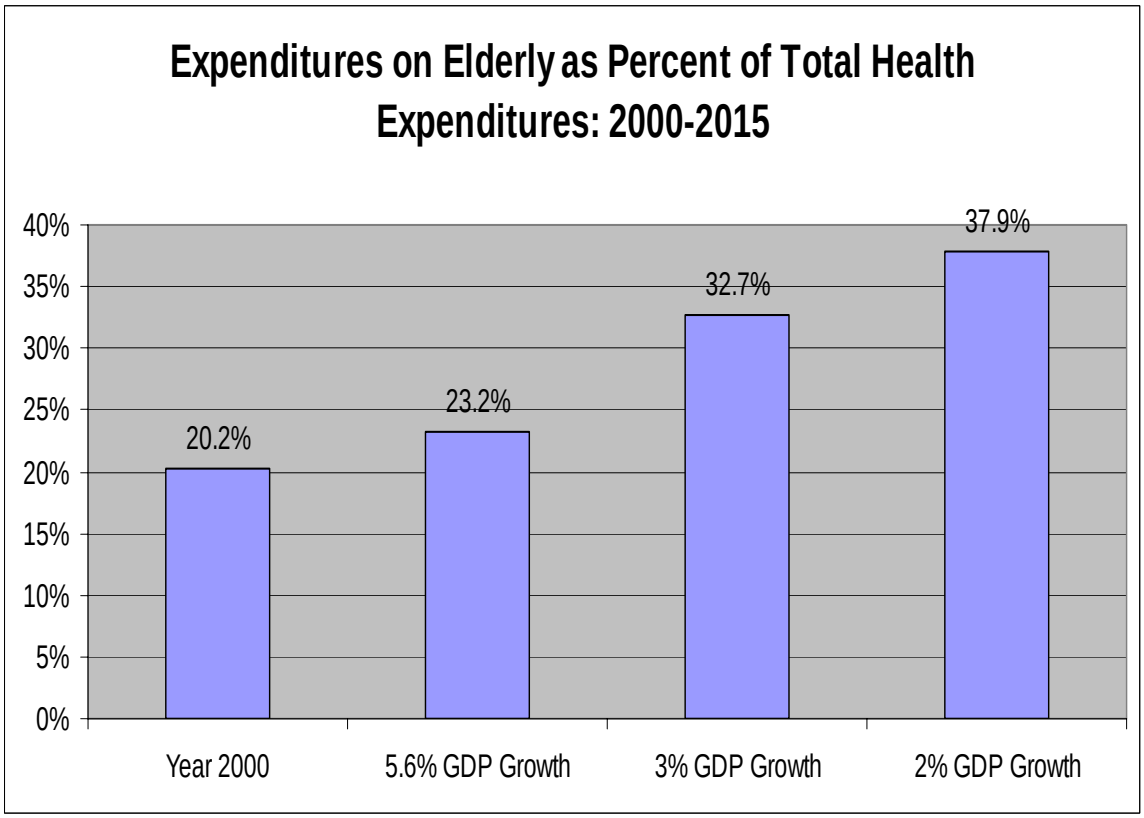
Similarly, globalization will force insurance companies to enter health markets in low and middle-income countries where longitudinal data traditionally used to price insurance products is not available. The methodology presented in this paper offers an alternative approach to estimating expenditures on the elderly in such situations. This method will also yield estimates of expenditures on other population sub-groups. As with all econometric projection models our approach has its limitations including assumptions that have been used to age the population. Similarly, the effects of sudden shocks to the system cannot be modeled in our approach. Even with these limitations the approach presented in this paper, with some country specific modifications, should be applicable in many low and middle-income countries.

Annex 1: Application of Methodology to Jordan






Percent Change in Annual Per Capita Out-of-Pocket Expenditures
Elderly:2000-2015
140%
120%
100%
80%
60%
40%
20%
0%
Outpatient
Inpatient
Medicines
5.6% Growth rate
3% Growth rate
2% Growth rate


Expenditures on Elderly as Percent of Total Health
Expenditures: 2000-2015
40%
35%
30%
25%
20%
15%
10%
5%
0%
20.2%
23.2%
32.7%
37.9%
Year 2000
5.6% GDP Growth
3% GDP Growth
2% GDP Growth