Centre for Health Economics and Policy Analysis

CHEPA WORKING PAPER SERIES

Paper 04-01

Developing Needs-based Funding Formulae Using Individual-level Linked Survey and Utilization Data: An Application to Home Care Services in Ontario, Canada

Jeremiah Hurley, PhD
Brian Hutchison, MD, MSc
Gioia Buckley, MA
Christel Woodward, PhD

McMaster University

NOT FOR CITATION WITHOUT PERMISSION

Developing Needs-based Funding Formulae Using Individual-level Linked Survey and Utilization Data: An Application to Home Care Services in Ontario, Canada*

Jeremiah Hurley, PhD[1,2,3]
Brian Hutchison, MD, MSc[2,3,4]
Gioia Buckley, MA[2,3]
Christel Woodward, PhD[2,3]

[1]Department of Economics, McMaster University
[2]Centre for Health Economics and Policy Analysis, McMaster University
[3]Department of Clinical Epidemiology and Biostatistics, McMaster University
[4]Department of Family Medicine, McMaster University

*Corresponding author: Jeremiah Hurley
Department of Economics - KTH 430
McMaster University
1280 Main Street West
Hamilton, Ontario L8S 4M4 Canada
Tel: 905-525-9140, ext. 24593
Fax: 905-521-8232
e-mail: hurley@mcmaster.ca

Cite as: Hurley J, Hutchison B, Buckley G, Woodward C. *Developing Needs-based Funding Formulae Using Individual-level Linked Survey and Utilization Data: An Application to Home Care Services in Ontario, Canada.* McMaster University Centre for Health Economics and Policy Analysis Working Paper 04-01. February 2004.

This research was supported by the grant from the Ontario Ministry of Health and Long-Term Care to the Centre for Health Economics and Policy Analysis. We thank a number of people who provided assistance in conducting this work, the Community Funding Review Committee, Vicki Rynard and Rob Hopkins. Thomas Crossley provided technical advice regarding model estimation. Paul Contoyannis, Greg Stoddart and members of the health Polinomics seminar group at McMaster University provided helpful comments. We retain responsibility for any errors.

Preface

The work presented in this paper derives from the same collaborative project between CHEPA and the Ontario Ministry of Health and Long-Term Care that generated two previous reports:

Hurley, J. B. Hutchison, G. Buckley and C. Woodward. 2002. *Needs-based Funding for Home Care and Community Support Services in Ontario: Final Report to the Community Funding Review Committee*. Ontario Ministry of Health and Long-Term Care. February.

Hurley, J., B. Hutchison, G. Buckley, and C. Woodward. 2003. *Needs-based Funding for Home Care and Community Support Services in Ontario: A New Approach Based on Linked Survey and Utilization Data.* McMaster University Centre for Health Economics and Policy Analysis Working Paper No. 03-01.

This preface clarifies the relationship between this paper and the above two documents. The goal of this paper differs from the two previous reports. Each of the above two reports document the process by which we, working with the Ontario Ministry of Health and Long-Term Care, developed a needs-based funding formula that could serve as the basis for allocating funding for home care and community support services in Ontario. The purpose of this paper is distinct. This paper emphasizes methodological and conceptual issues encountered in developing needs-based funding formulae, and is written in a style appropriate for an academic journal. The empirical models presented in this paper differ slightly from those reported in the previous reports (in general they are more parsimonious), largely because the purpose is to illustrate conceptual and methodological issues rather than report the "best" point estimates for each region that could serve as the basis for MOHLTC funding. It is inappropriate to draw any conclusions regarding appropriate funding levels for CCAC regions by comparing the results from this paper with previous reports. As a precaution, we have therefore changed the CCAC coding scheme when presenting results in this paper.

Abstract

A common goal of health policy is to allocate public health care resources according to need. This paper presents a method for developing needs-based funding formulae using individual-level linked health survey and utilization data. Needs-based funding shares are developed in three basic stages: (1) estimate the full utilization model, including both need-related and non-need-related adjustors; (2) predict individual-level needs-based home care utilization holding all non-need factors constant; (3) use individual-level estimates and sample weights to develop regional needs-based allocations. The results suggest that methods based on such data offer considerable potential; however, they also raise several new challenges.

1. Introduction

Smith, Rice and Carr-Hill (2001) argue that, because need-based funding models based on regional-level relationships suffer from a number of inherent limitations (e.g., potential for ecological fallacy), advances in such funding formulae in health care will come from employing micro-data (Smith et al 2001). Individual-level data are commonly used in the development of risk-adjusted capitation within competitive (usually) private health insurance markets (van de Ven & Ellis 2000).[1] Data limitations, however, mean that such approaches often can incorporate only a limited set of individual characteristics such as age, sex, an indicator of health status (often derived from utilization-based diagnostic information) and, perhaps most importantly, prior health care utilization. Prior utilization is the best single predictor of current-period utilization. But because a key objective of needs-based funding is to redress historical inequities of access and utilization, many argue that prior utilization is an inappropriate adjustor for needs-based funding formulae. Hence, beyond information on age and sex, micro-data have played a limited role in needs-based funding.

In this paper we present an approach to the development of a needs-based capitation funding formula that employs a unique micro-data set linking, at the individual-level, population-based health-survey information with utilization data drawn from administrative databases of the universal, public insurer. The data provide detailed information on individuals, their household and their health care utilization. We use the data to develop a needs-based funding formula to allocate the public home care budget in the Province of Ontario, Canada, which contains 43 geographically defined Community Care Access Centre (CCAC) regions responsible for managing the provision of home care services to residents within each region. Individual-level models derived from such data offer considerable potential for the development of needs-based funding formulae. Formulae based on such data capture a greater proportion of variation in health care use and needs while avoiding ecological problems associated with regional-level data. Micro-data, however, also raise a number of issues not heretofore emphasized in the needs-based funding literature, including endogenous individual-level relationships, classification of variables (particularly certain demand-side variables) as need- or non-need-related, and validation strategies.

In the next section we provide a brief discussion of the underlying motivation for allocation by needs-based formula; section 3 describes our data and empirical strategy; section 4 presents our conceptual model of home care utilization; section 5 presents our empirical model; section 6 presents results; section 7 discusses implementation; and section 8 offers conclusions.

---

[1] We distinguish risk-adjustment from needs-adjustment. The goal of risk adjustment is actuarial — to predict actual utilization. The goal of needs-adjustment is to predict need, which can differ from utilization because of both unmet needs and over-utilization. The two tasks therefore assess a funding formula against different criteria. They increasingly, however, use similar data and related methods.

2. Allocation According to Relative Need

Allocation of health care resources according to need is a fundamental, explicitly identified health policy goal in a number of countries (Rice & Smith 2001, Wagstaff & and van Doorslaer 1993).[2] Central governments increasingly implement this policy goal through needs-based capitation formulae that allocate health care resources to geographically defined regional authorities in accordance with the relative need for health care among regions (i.e., a region with 10% more needs should get a 10% larger share of the central budget) (Smith et al 2001).

Four economic considerations generally underlie allocation by explicit formulae. First, even with substantial public subsidy, often to the point of zero price, aggregate demand in some regions is insufficient to attract health care providers, so the forthcoming supply is inadequate to meet needs. Second, informational problems associated both with diagnosing illness and with identifying effective health care interventions conditional on diagnosis mean that individuals are often unable to judge their need and base demands on accurate assessments of need. Third, these same informational problems confer considerable market power on providers, which can generate a distribution of health care providers that does not correspond well to the distribution of health care needs. It is frequently the case that urban areas are judged to have an oversupply of providers and rural and remote areas an undersupply. Fourth, much health care infrastructure in many countries was developed when private financing played a dominant role in health care finance, and private finance may still play an important role in mixed systems of finance. The distribution of facilities and providers, therefore, reflects systematic geographic variation in ability-to-pay as well as in need. The importance of each of these varies by setting and health care sector, but together they imply that current allocations, even within universal, first-dollar, publicly financed health care systems, may depart substantially from allocation according to need. Allocation to regional authorities according to an explicit needs-based formula has the potential to redress, at least in part, such departures from need.

Arguments in support of the relative-need principle, which calls for resources to be allocated in accord with the *relative* need for health care among regions, have chiefly been pragmatic. The relative-need principle allows one to divide the allocation problem into two components each of which requires distinct information and expertise: (1) determination of the size of the budget, which is a political decision; and (2) division of that budget (whatever its size) among regions on the basis of relative need, which is framed as a technical problem for which analysts have much to contribute. In addition, however, limited empirical evidence on individual preferences for the allocation of health care resources supports the relative-allocation rule. In a series of allocation experiments, both Yarri and Bar-Hillel (1984) and Kahneman and Varey (1991) found that when asked to allocate a fixed quantity of a health producing good whose supply was insufficient to meet total need among a set of individuals, a sizeable majority chose the allocation that equalized the proportion of met needs across individuals, i.e., the relative-need rule. See Hurley (2003) for a model which derives conditions under which allocation according to the relative-need principle is optimal within an economic framework of utility-maximizing individuals, defined health production technology, and a social welfare function aggregated from individual utilities.

3. Data and Empirical Strategy

Our objective was to develop a needs-based capitation formula for home care services that would allocate a given provincial home care budget in line with relative need for home care across the 43 home care regions in Ontario, Canada. Our data included detailed information on personal characteristics and health care utilization for a representative sample of the Ontario population. The strategy for developing the needs-based capitation rates for public home care funds included six key steps:

1. *Estimate each sample respondent's expected utilization of home care services (EHCU)*. We measure utilization in terms of the dollar value of home care services received. Utilization is determined by both need-related factors, $\mathbf{X}^{\mathbf{N}}$, which from the perspective of public policy should influence a needs-based allocation, and non-need-related factors, $\mathbf{X}^{\mathbf{NN}}$, which should not influence a needs-based allocation):

$$EHCU_i = \beta X_i^N + \delta X_i^{NN} + \varepsilon_i \tag{3}$$

2. *Estimate each sample respondent's expected needs-based home care utilization (ENBHCU)*. To predict needs-based utilization, we use (3) above but, for each element of X_i^{NN}, replace individuals' values of the variable with the sample mean for the variable:

$$ENBHCU_i = \hat{\beta} X_i^N + \hat{\delta} \overline{X}^{NN} \tag{4}$$

[2]The recent UK policy directive to allocate funds to reduce avoidable health inequalities is a notable exception to this pattern. Hauck (Hauck et al 2001) discusses the implications of this alternative objective for resource allocation.

Substituting the sample mean of non-need factors ensure that during prediction there is no variation in X^{NN} across the sample, so these factors do not influence the needs-based allocation across regions.

3. *Estimate the need for home care in each region ($NBHCU_r$).* To estimate each region's aggregate home care need, inflate each respondent's expected needs-based home care utilization using the sampling weights and, within each region, sum across respondents from the region.

$$NBHCU_r = \sum_{i \in r} ENBHCU_i \cdot WGHT_i \qquad (5)$$

4. *Estimate the aggregate, provincial home care need ($NBHCU_p$).* To estimate the total provincial need, sum across the regional need estimates obtained in (5).

$$NBHCU_p = \sum_{r=1}^{R} NBHCU_r \qquad (6)$$

5. *Estimate each region's relative needs-based share (NBS_r).* To estimate each region's relative needs-based share, divide each region's aggregate need into the provincial aggregate need:

$$NBS_r = \frac{NBHCU_r}{NBHCU_p} \qquad (7)$$

6. *Estimate each region's needs-based capitation rate ($NBCR_r$).* To obtain each region's needs-based capitation rate under a given provincial budget, multiply each regional share by the provincial budget and divide by the regional population:

$$NBCR_r = (NBS_r \cdot BUD_p) / POP_r \qquad (8)$$

Assuming the validity of the sample and the associated population weights, the empirical validity of the approach rests on the first two steps -- our ability to estimate expected home care utilization and then purge expected utilization of the influence of non-need drivers of utilization. We therefore describe in detail below our underlying conceptual model of home care utilization, our classification of variables as need- or non-need-related, and our empirical estimation strategy.

4. A Conceptual Model of Home Care Utilization

4.1 Demand for Home Care

We posit the following demand for publicly financed home care for individual i in region r:

$$D_{ir} = d(PN_i, P_r^{pub}, P_r^{priv}, INC_i, PREF_i, ADV_i, RSRV_r) \qquad (9)$$

where:

D_{ir} = quantity of public home care demand by individual i in region r

PN_i = perceived need for home care by individual i

P_r^{pub} = price of public home care services

P_r^{priv} = price of private home care services

INC_i = income of individual i

$PREF_i$ = individual i's preferences with respect to home care

ADV_i = advocacy ability of individual i in obtaining home care services

$RSRV_r$ = terms and availability of related services (e.g., long-term care facilities)

Prices, income, and availability of related services are standard determinants of demand. Preferences for home care are intended to represent the desire for home care services conditional on a level of perceived need: a person may want care that even they do not perceive as "needed", just as they may not want care that they perceive to be needed (e.g., because they do not like having "strangers" in the house). Advocacy ability (ADV_i) refers to a person's ability to navigate the care-seeking process and assert one's demand within a complex, supply-constrained health care system. It depends on the individual's knowledge of the system (KS_i), skills in navigating the system (SK_i) and on the availability of others such as family members who advocate on behalf of the individual (AGT_i).

$$ADV_i = a(KS_i, SK_i, AGT_i) \tag{10}$$

Perceived need for home care (PN_i) is a function of true underlying need (N_i), as well as one's attitudes toward and knowledge of one's health condition, potentially effective services and what constitutes a health care need (AK_i).

$$PN_i = m(N_i, AK_i) \tag{11}$$

Underlying need (the focus of needs-based funding) is a function of health status (HS_i), the level of informal support and care available from family members, neighbours, and friends ($ISPRT_i$), and socio-economic factors (SES_i) (which can affect need conditional on health status):

$$N_i = n(HS_i, ISPRT_i, SES_i) \tag{12}$$

Finally, the attitudes and knowledge that affect perceived need (AK_i) are shaped by a range of factors, including socio-economic status (SES_i), ethnicity (ETH_i), previous contact with the health care system ($PREV_i$), and provider influences ($PROV_i$).

$$AK_i = k(SES_i, ETH_i, PREV_i, PROV_i) \tag{13}$$

Substituting (10) - (13) into (9), we get:

$$D_{ir} = d(P_r^{pub}, P_r^{priv}, INC_i, PREF_i, RSRV_r, HS_i, ISPRT_i, SES_i, KS_i, SK_i, AGT_i, ETH_i, PREV_i, PROV_i) \quad (14)$$

4.2 Supply of Home Care Services

Each CCAC region in Ontario receives a budget to purchase home care services for residents through an internal market. The supply of home care services in CCAC region r, S_r, is therefore determined by the size of the region's budget (B_r) and the prices it must pay for the various services it purchases ($PSRV_r$):

$$S_r = s(B_r, PSRV_r) \quad (15)$$

The price a CCAC region must pay for services depends on the prices of inputs into service production (INP_r) and the CCAC's skill in negotiating and managing contracts with provider agencies ($MGMT_r$):

$$PSRV_r = c(INP_r, MGMT_r) \quad (16)$$

The home care system is supply-constrained with no price adjustment to equilibrate the market. Rather, services are rationed through needs-assessment by case managers and the associated queues that result. Both the likelihood that an individual will utilize services and the quantity of services utilized, are therefore a function not of aggregate supply, but of aggregate supply adjusted for aggregate need/demand in the region. Other things equal, a person residing in a region that has historically been funded above need is more likely to utilize home care than is an identical person residing in a region that has been underfunded relative to need. Therefore, we define adjusted supply as:

$$AS_r = t(INP_r, MGMT_r, B_r, AN_r) \quad (17)$$

where AN_r is a measure of aggregate need.

4.3 Utilization of Home Care Services

Observed utilization results from the interaction between adjusted supply and demand, though in this context no price mechanism equilibrates aggregate supply and demand in each region:

$$U_i = u(P_r^{pub}, P_r^{priv}, INC_i, PREF_i, RSRV_{r,}, HS_i, ISPRT_i, SES_i, KS_i,$$
$$SK_i, AGT_i, ETH_i, PREV_i, PROV_i, INP_r, MGMT_r, B_r, AN_r) \quad (18)$$

Equation 18 highlights a number of challenges associated with empirically developing needs-based funding using individual-level utilization data.

First, the goal of the analysis is to identify true need (Eq. 12). However, we observe utilization, which measures true need with error. If the measurement error is random the only consequence is a loss of efficiency in estimation. But correlation between the measurement error and any of the adjustment factors included in the empirical model introduces bias. The potential for such bias must be a concern given that a prime motivation for needs-based funding is the judgement that under- and over-utilization are systematic rather than random. Hence, it is crucial to model empirically the utilization process as fully as possible to reduce possible sources of correlation between the error term and the adjustment factors.

Second, purging utilization of non-need-related components requires that we classify each determinant of utilization as either need-related or non-need related. Schokkaert and Van de Voorde (2003) highlight a directly analogous point in a recent paper on health care funding models. The needs-based literature to date has emphasized almost exclusively supply-side factors as non-need drivers of utilization (Smith et al 2001). Modelling individual-level utilization, however, highlights the potential for demand-side factors to generate non-needed utilization. The classification of some of these demand-side factors is more difficult, more controversial and likely more context-specific than is the case for supply-side factors. [3] What level of support, for instance, should be expected from a relative before publicly financed formal services are needed? Should wealthier regions get fewer public resources because they have greater access to private care? How should attitudes be treated? These questions must be confronted directly when deriving needs-based funding formulae using individual-level data.

Third, a number of factors influence utilization through both need-related and non-need- related paths. Socio-economic status, for instance, simultaneously affects utilization through its effects on true need, on attitudes regarding perceived need and on preferences for home care services. The data are insufficient to estimate a structural model. For such variables we can estimate only the reduced-form parameters that reflect both need and non-need pathways. The ultimate classification of such factors as need or non-need at the prediction stage is therefore contingent on which effect dominates empirically.

[3] See, for example, Mooney *et al.* (Mooney et al 1991, Mooney et al 1992) and Culyer *et al.* (1992b, 1992a), who debated the treatment of demand-side factors in economic analysis of the distribution of health care resources.

5. The Empirical Model of Home Care Utilization

5.1 Two-part Utilization Model

The vast majority of individuals do not use home care and, among those who do, the distribution of expenditures is highly skewed. We therefore employ the two-part utilization model (Jones 2000):

$$E(hc|x) = \quad Pr(hc > 0 \mid x) \bullet E(hc \mid hc > 0, x) \tag{19}$$

where: hc = home care expenditure

x = a vector of individual and regional-level variables that influence utilization of home care

The first part, Pr(hc > 0| x), is estimated over the full sample (N=22,855) via logistic regression with a dichotomous dependent variable indicating whether an individual used any home care services. The second part of the model, E(hc | hc > 0, x), is estimated over only the sub-sample of home care users (N = 1447), employing either a log-OLS specification or a General Linear Model (GLM) (Duan 1983, Manning & Mullahy 2001, Mullahy 1998, Manning 1998, Blough et al 1999). We followed Manning and Mullahy's (2001) recommendations in choosing between log-OLS and GLM specifications.

5.2 Data and Variable Specification

Data for the study come from the Ontario component of Canada's 1995/96 National Population Health Survey, which was designed to collect information related to the health of non-institutionalized Ontarians and which had sufficient sample size to provide reliable cross-sectional estimates at the sub-provincial, health-area level (Statistics Canada 2002).[4] Health-survey information of particular relevance to this study included an individual's age, sex, marital status, self-assessed health status, chronic conditions, activities of daily living requiring assistance, functional status, social support, living arrangements, household income, education level, and ethnicity. These survey data were linked to administrative data from the Ontario Ministry of Health and Long-Term Care regarding individual-level utilization of direct home care services, general practitioner services and inpatient hospital admissions (overnight and day procedures).

[4]There are 23 health districts in Ontario and 43 CCAC regions. Hence, we need to make inferences at a finer level of disaggregation than planned for in the survey design. We return to this below in the discussion of the precision of our estimates. No CCAC regions cross health district borders; 12 of the 23 health districts fully coincide with CCAC regions; 6 health districts each contain 2 CCAC regions, 3 health districts each contain 3 CCAC regions, 1 health district contains 4 CCAC regions and 1 health district (Toronto) contains 6 CCAC regions.

5.3 Dependent Variables

Utilization is defined as the dollar value of direct, publicly funded home care services received during the 32-month period surrounding the respondent's survey interview date (i.e., 16 months before and 16 months after the interview date).[5] The dichotomous dependent variable for part 1 indicates whether a respondent used any publicly financed home care services in the 32-month period; the dependent variable for part 2 is the dollar value of publicly financed home care services received in the 32-month period.

5.4 Independent Variables

Table 1 lists the variables included in the model and briefly describes the rationale for their inclusion. With the exception of the price of private home care services, we are able to include at least one variable representing each of the key factors in our model of home care utilization.[6]

A number of the variables likely capture multiple influences. Education and income, for example, are intended respectively to represent knowledge and ability-to-pay for private health care. But both may also capture general socio-economic influences and, given the well- established correlation between health status and socio-economic status, unmeasured aspects of health status. The number of General Practitioner/Family Physician (GP/FP) visits is included to represent unmeasured aspects of health status, but may also represent an individual's general propensity to use health care and the influence of a provider advocating on behalf of the individual to obtain home care.

Some variables are potentially endogenous, particularly health status, functional status, income, living arrangement, GP/FP visits, and hospitalization. Reverse causation from receipt of home care to health status, income and living arrangement is unlikely to be empirically important. Home care use arises in two primary contexts: following an acute inpatient procedure and to support frail elderly. In the former, home care substitutes for care that otherwise would have occurred in hospital and is unlikely to generate long-term health effects independent of the inpatient procedure performed. In the latter, most home care consists of personal support services, not health care services *per se*. In either

[5] The expenditures exclude case management and overhead costs, which are not allocated to individual clients. The 32-month period was determined in consultation with members of the Ministry's Community Funding Review Committee, which included policy makers and stakeholders from the home care sector. The cut-off balanced the desire for a sufficient number of users against the decreasing strength of association between characteristics measured at the interview and home care use, as the potential time difference between them increases.

[6] We did not have access to price information. We did, however, have data on region-specific average cost per unit of service (by category of service). An auxiliary analysis found no relation between average cost per service and a region's population density, total population and relative funding level.

case, home care is unlikely to have a large independent effect on health status.[7] The causal link from home care to income is likely to be weak because the largest users of home care, the elderly, depend primarily on unearned income. Finally, although home care likely affects the probability that an individual can continue to live alone, its availability is unlikely to induce someone to change from living with someone to living alone (which would be a problem for our analysis).

Endogeneity may be more serious for hospital admissions and GP/FP visits. Two potentially offsetting effects operate for each. Receipt of home care may decrease the rate of hospitalizations (GP visits) by dealing effectively with health problems before they become serious. But receipt of home care may increase the rate of hospitalizations (GP visits) if a home care professional identifies needs that otherwise would have gone unrecognized and unmet. Although the net effect is unknown for GP visits, the former likely dominates in the case of hospital admissions: other things equal, home care should reduce hospitalizations, so any bias should lead to a conservative estimate of the effect of hospitalization on home care use.

The hospitalization variable may also capture a spurious effect. Hospitalization is included primarily to represent short-term need for home care following an acute-care hospital episode. In some cases, however, lack of home care may cause failure in the home, leading to a hospitalization, at which point the need for home care is identified. To the extent this occurs, the empirical association between home care and hospitalization will be spurious. The magnitude of this effect can be assessed using a lagged hospitalization variable. Because home care use following an acute care episode is generally short-term, such utilization should generate no relationship between previous-period hospitalization and current-period home care use. In contrast, home care need following hospitalization caused by gradual decline is chronic, generating a strong relationship between previous-period hospitalization and current-period home care use.

5.5 Need vs Non-Need Variables

Needs-adjustment requires that we must classify each variable as need-related or non-need-related. Table 1 lists our classification of each variable as definitely need-related (N), definitely non-need-related (NN) or contingent (C), meaning that its classification depended on the sign of the estimated empirical relationship. We judged that the following factors definitely should influence a needs-based allocation: age, sex, self-assessed health status, number of chronic conditions, number of activities of daily living requiring assistance (ADL), hospitalizations, GP/FP visits, marital status, living arrangements, social support, contact with neighbors, income. The link between most of these and

[7] If receipt of health care is perceived by an individual to indicate lower health status, a person receiving formal home care might, other things equal, rate their health status to be lower than someone not receiving

need for home care is clear. GP/FP visits are classified as need-related because of the potential association between GP/FP visits and unmeasured aspects of health status. Income is classified as need-related because its relation to both unmeasured health status and ability-to-pay lead us to expect that, other things equal, those of lower income will utilize more public home care services, so the allocation should reflect these influences. We should emphasize that, by classifying variables associated with informal support (e.g., marital status, living arrangements, etc.) as need-related, we imply that need for publicly financed formal services arises only after such informal support becomes inadequate.

The following factors are classified as definitely not need-related: historical adjusted CCAC budget, English language ability, acute-care-induced-home-care days per capita, and CCAC-region dummy variables. We expect a positive association between adjusted historical funding and utilization and do not want previous over/under-funding to influence a needs-based allocation. Similarly, acute-care-induced-home-care days per capita is a measure of aggregate demand placed on home care resources in the region by activity in the acute-care hospital. Other things equal, we expect this to have negative effect on an individual's utilization, an effect that should not affect a needs-based allocation. Inability to speak English represents a potential barrier to utilization that should not influence a needs-based allocation. The regional dummy variables capture a host of unmeasured features of a region that are associated with utilization of home care by residents of the CCAC region. If, conditional on all the factors included in the model, residents of a CCAC region systematically have lower (higher) than expected utilization, we interpret this to represent unmet need (over-utilization). Such effects should not influence a needs-based allocation.[8]

The classifications of the following variables are contingent because of countervailing forces. To the extent that education level captures knowledge and ability to navigate the system, we expect a positive association with home care use; to the extent it captures unmeasured health status, we expect a negative association. If the former dominates, education should be classified as a non-need factor; if the latter dominates, it should be classified as a need-related factor. Aboriginals have generally higher levels of need but, as a marginalized group in Canadian society, they often face barriers to care, lowering utilization. Both long-term care beds and physicians may substitute for home care, generating a negative relationship with home care use (and suggesting a need-related classification); but if their level of supply is indicative of the more general supply of health care in a region, we may observe a positive correlation, in which case they should be classified as non-need.

---

home care.

[8] This interpretation is not inconsistent with the source of such variation being systematic differences (ceteris paribus) in preferences for home care across regions. But, in accord with the extra-welfarist orientation of needs-based funding, such preferences are subordinated to need-related factors.

6. Results

We started with a broad, comprehensive specification of the model, and then eliminated variables using explicit statistical criteria relating to individual and/or joint significance tests as well as a range of specification tests (e.g., omitted variables, heteroskedasticity, normality of residuals, etc.) to arrive at a relatively parsimonious model that retains substantial explanatory power and meets basic statistical criteria. We in general specified variables to allow for non-linear associations with utilization and we tested the following interactions: age interacted with all other variables, sex interacted with all other variables, and health status interacted with selected demographic (living arrangements, social support) and socio-economic (education, household income per capita) variables. All models were estimated in STATA using weighted analysis; standard errors reflect adjustment for the complex clustered sample design (Statistics Canada 2002). Table 2 presents descriptive statistics on the variables.

6.1 Part 1: Use/Non-Use

Table 3 presents the results of the weighted logistic regression analysis of use/non-use. It includes only those factors that met our statistical criteria for inclusion. Compared to a model with only age and sex terms (the only factors in Ontario's current capitation formula), the final model explains substantially more variation in the use of home care (pseudo-R^2 of 0.41 vs. 0.19), has smaller prediction bias (0.0003 vs. -0.0049), smaller mean squared prediction error (0.04 vs. 0.05), and is better able to distinguish users and non-users (difference in mean predicted probability of 0.34 vs. 0.15).

The identified associations with home care utilization are, in general, as one would expect. Age is the dominant determinant of home care use, both directly and indirectly in interaction with other factors in the model. The probability of home care use rises non-linearly over the entire life cycle. There are important interaction effects between age and a number of other factors; such effects are particularly strong between age and activities of daily living, age and living status, and age and hospitalization. Because such interaction effects complicate assessment of the impact of each factor, the bottom section of the table lists the change in probability of home care use at different points in the life cycle given differing levels (or categories) of each factor.[9] The two most important determinants of home care use besides age, and the only two factors that exert quantitatively important influence over the entire life-cycle, are the number of activities of daily living for which assistance is required

[9] The base case is male, unmarried, living alone, in good health, with no chronic conditions, requiring no assistance with ADLs, no hospitalizations, with secondary diploma, and income in the 3rd quartile.

(representing chronic need), and the number of hospitalizations (representing acute need).[10] In both cases there is the expected gradient across levels of these factors: at age 90 for example, compared to a person who requires no assistance, the probability of home care use increases by 0.104, 0.139, 0.281, 0.442 and 0.460 for persons who require assistance with 1, 2, 3, 4, and 5 ADLs respectively; similarly, among those aged 90 the probability of home care use is 0.198 and 0.312 higher for someone with 1 and 2 or more hospital admissions respectively compared to someone with no hospitalizations.

Health status exhibits the expected gradient with utilization, though it is not quantitatively important except perhaps for elderly who rate their health as poor (e.g., probability of home care use is 0.05 higher for a person aged 90 in poor health than for the same person in excellent health). Important differences by sex, living arrangements, education level, and income emerge only late in life, at which point females have a higher probability of use than males, those living alone have a higher probability than those living with someone, and those with less educational attainment and lower income have a higher probability of use.

CCAC-region dummies (not shown) were retained for eight CCAC regions for which, conditional on all other factors, utilization by residents was statistically different from the provincial average level of utilization (Suits 1984, Kennedy 1986). The coefficients on individual-specific variables were insensitive (i.e., change by less than 5%) to inclusion or exclusion of region dummies, indicating no important correlation between unobserved regional-level factors and individual-level adjustors. Not surprisingly, region-specific variables were more sensitive to inclusion/exclusion of region dummies. Long-term care beds-per-capita and historical funding were significant when region dummies were excluded, but they became non-significant when region dummies were included.

6.2 Part 2: Predicting Home Care Expenditures Among Home Care Users

The procedures outlined in Manning and Mullahy (2001) indicated mild to moderate heteroskedasticity under the logged-OLS specification with an error structure that would be difficult to correct for during re-transformation. The analysis further indicated that a GLM model with a gamma link function was most appropriate. Therefore, a weighted gamma-GLM model with a log link was used to estimate expenditures conditional on being a home care user (Table 4). Once again, the full model explains substantially more of the variation in home care expenditures than does a model based on age and sex alone. The strength of the association between the potential adjustors and expenditures, however,

[10]When lagged rather than current hospitalization was specified, the estimated coefficient was either non-significant or marginally significant and was only one-fifth the size of the relationship for the concurrent hospitalization variable. Hence, although some of the estimated effect may represent a spurious association, the vast majority represents acute, post-hospital need.

is not as strong as between adjustors and the probability of home care use. Age, for instance, does not have any independent association with expenditures; it enters the model only in interaction terms with other variables. Once again, the most important factors are the number of activities of daily living for which assistance is required and the number of hospitalizations. In the former case, there is again a strong gradient over the entire life cycle. In the latter, up to age 78 those who have a single hospitalization have lower costs than those who are not hospitalized, and this difference decreases over the life-cycle. This is plausible as hospitalized individuals require home care for only a short period following discharge, while non-hospitalized users likely suffer from chronic problems requiring long-term home care. However, those who are hospitalized two or more times have higher expenditures over almost the entire life-cycle, suggesting that the variable may be picking up some unmeasured aspects of health status in addition to any acute home-care needs. The expenditures patterns with respect to chronic conditions and sex exhibit a cross-over. Below age 66, females have lower expenditures than males but the opposite is true above age 66. Similarly, below age 72, those who suffer from fewer than four chronic conditions have higher expected expenditures than those with four or more, while the opposite is true above age 72. We can offer no plausible explanation for this unexpected pattern. Finally, those with less than grade 9 education have lower expected expenditures than those who attained grade 9 or higher, suggesting that either a barrier effect dominates among those with low education levels or that they have different needs and use a different, lower-cost mix of services.

The inclusion/exclusion of CCAC-region dummies had larger effects on coefficient estimates than they did in part 1, though the pattern of statistical significance on other variables remained unchanged by their inclusion/exclusion.

6.3 CCAC Regional Capitation Rates

An individual's expected home care expenditure is simply the product of her predicted probability of use and predicted expenditures conditional on being a user. (This is also referred to as the "risk-adjusted" expenditure.) To estimate a person's expected needs-based home care expenditure, we predict expenditure setting the values of all non-needs-related factors equal to their sample means (as per (4) above). Regional capitation rates were then derived as described in equations (5) through (8) above.

Table 5 presents the CCAC-regional needs-adjusted, actual and risk-adjusted capitation rates (expressed in dollars per capita) based on the 1999-2000 provincial budget. Confidence intervals for the needs-adjusted and risk-adjusted rates were estimated using a bootstrap procedure (Statistics Canada 2002). The needs-adjusted capitation rates vary nearly seven-fold ($55 – $347) across regions, which is substantially larger than variation in actual rates ($69-$167). Actual funding exceeds the needs-based estimate by more than 30% for 5 CCAC regions, more than 20% for 11 CCAC regions and more than 10% for 17 CCAC regions; actual funding falls below needs-based funding by more than 30% for 6 CCAC regions, more than 20% for 9 CCAC regions. Because home care use is a relatively rare event in the general population, the standard errors and confidence intervals reveal considerable imprecision in the needs-based estimates for some CCAC regions. At conventional levels of statistical significance we can conclude that a true difference exists in 6 CCAC regions (actual exceeds needs-based in 7 cases; actual is below need in 1 cases). We performed an auxiliary regression analysis to examine the association between the difference in actual and needs-based capitation rates and regional characteristics (e.g., proportion in excellent or very good health, age distribution, supply of physicians, etc.). The sample is small (N = 43), so power is weak, but the analysis found no systematic relationship between the difference in capitation rates and regional characteristics.

The large difference in point estimates between the needs-adjusted and risk-adjusted methods indicate the importance of needs-adjustment (i.e., purging estimates of the influence of non-need drivers of utilization), though once again the large standard errors prevent us from concluding that a true difference exists. The most important non-need risk factor purged in the process of estimating expected needs is the region-specific intercept, which captures a range of unobservable factors in a region. As one would expect, there is a positive correlation between needs-adjusted and risk-adjusted rates ($\rho = 0.68$). However, there is no systematic relationship between the level of need and the direction or size of the difference between need and risk-adjusted rates.

7. Implementation

The optimal allocation can be achieved through redistribution by taking funds from over-funded regions and reallocating to under-funded regions. In practice, governments in most jurisdictions have been unwilling to reduce the absolute level of funding to any region, so that the needs-based funding formula guides only the allocation of new funding, often based on *ad hoc* incremental adjustment rules. From an economic perspective, however, incremental funds should be allocated so as to obtain the greatest increase in social welfare. Hurley (2003) shows that, based on the empirical approach used to estimate the needs-based capitation rates and the specification of the social welfare function for which the relative-need allocation rule is optimal, the first derivative of the social welfare function can be estimated (up to a factor of proportionality) by the statistic, $-\frac{HC_r}{HC_r^*}$, where HC_t is the actual capitation rate for CCAC region r and HC_r^* is the needs-based capitation rate for region r. It is also possible to derive confidence intervals around the point estimate of this statistic. We call $\frac{HC_r}{HC_r^*}$ the "derivative statistic." The region with the smallest derivative statistic evaluated at its current funding level (i.e., the largest derivative) has the greatest priority for new funds (See Appendix 1 for additional details).

Table 6 lists these point estimates and standard errors for the derivative criterion for the 43 home care regions in Ontario based on their 1999-00 actual budget levels, as well the capitation rates that would result after the optimal allocation of a 5% budget increase. The budget increase is allocated by first giving additional funding to the region with the smallest derivative statistic until its value, evaluated at its new funding level, equals the derivative statistic of the next-lowest region, after which incremental dollars are allocated equally to both regions. The process continues until the incremental funds are exhausted. For a 5% budget increase, 11 of the 17 under-funded regions obtain additional funding.

8. Discussion

Although the specific motivation for this research was the desire to develop a needs-based funding formula for home care services in Ontario, Canada, the analysis makes a number of contributions to the needs-based funding literature. The results demonstrate the power of individual-level, linked health survey and health care utilization data to represent the relationship between utilization and its determinants. The number and detailed nature of the adjustors available exceed those used previously in developing needs-based funding models. More important from a methodological perspective, however, the individual-level models revealed non-linearities in the relationship between utilization and model adjustors -- non-linear models were required and interaction terms were very important -- implying that regional-level need depends on the full distribution of these variables. Hence, even if one could obtain analogous regional-level information (e.g. means, medians, proportions, etc.) on an equally rich set of adjustors, a regional-level model could not accurately represent regional-level need. Although survey and administrative data linked at the individual level through a common unique identifier are not readily available in many jurisdictions, in many settings it may be possible to achieve high-quality linkages. If sufficient personal identifying information is contained in both health survey and administrative data files, it may be possible to use probabilistic matching techniques to link the survey and administrative data with acceptable degrees of accuracy. Where such surveys are conducted only infrequently, impinging on the ability to update models, it is also possible to integrate models estimated using such linked survey data with regularly updated population-based information on, for instance, age-sex adjusted utilization rates. One such integrated approach proceeds as follows: (1) perform the usual age-sex standardization based on age-sex distribution of the population data and age-sex specific utilization rates; (2) use the linked survey data to estimate a model with only age and sex adjustors and calculate regional capitation rates based on this model; (3) use the survey data to estimate a full model that adjusts for age, sex and other factors and calculate regional capitation rates based on this full model; (4) calculate the percentage difference between (2) and (3); and (5) apply this percentage adjustment to the population-based age-sex standardized rates calculated in (1). This approach allows the age-sex standardized rates and adjustment for population distribution to be updated regularly, while the age-sex and full models are updated less frequently as survey data are available. The key point is that creative use of data may allow one to improve on simple regional-level models even if one can move only partially toward comprehensive data linked at the individual-level.

Approaches based on linked utilization/survey data can also incorporate diagnosis-based case-mix approaches such as Hierarchical Condition Categories (HCCs) and Adjusted Clinical Groups (ACGs), which are gaining popularity in the risk-adjustment literature (van de Ven & Ellis 2000). It may also be advantageous to combine the individual-level information with a more fully specified set of regional-level indicators within a multi-level modeling framework, potentially combining the advantages of individual- and regional-level modeling approaches.

Our results highlight the importance of calculating the precision of the regional needs-based estimates, an analytic feature heretofore largely absent from the needs-based literature. This proved particularly important in this application because home care is a rare event in the general population, so that needs-adjusted capitation rates could only be estimated with a relatively large degree of imprecision.

Approaches based on micro-data also raise a number of challenges that future work will have to address. As noted above, individual-level needs-based models require difficult judgments to classify each factor, and in particular demand-side factors, as need-related or non-need-related. As emphasized in A. William's seminal analysis of need, necessary conditions for a need to exist are: (a) that a service have positive marginal product with respect to a stated objective; and (b) the objective be seen a socially legitimate such that it warrants classification as a need rather than a want (Williams 1978). When allocating public health care funds, the pivotal issue is whether a particular characteristic of an individual or his/her situation deems it appropriate that s/he draw on public resources. Some factors are uncontroversial in this respect and would likely be universally classified as a need-related adjustor, e.g., age. Others are uncontroversial and would likely be universally classified as non-need-related, e.g., the pure effect of supply-side adjustors. But a wide range of factors, especially those under the control of the individual and closely linked to individual preferences, fall into a grey zone. Different jurisdictions might reasonably classify differently (as what is considered socially legitimate differs across cultures), and very often the classification of such factors will be viewed by some as controversial. What amount of assistance is a family member expected to provide, for instance, before the load becomes sufficiently high to warrant formal home care services at public expense? Similarly, should an individual who lives alone be eligible for a greater amount of subsidy than an identical individual who chooses to live with another individual? In this case, the positive marginal benefit of formal home care is due not to health status *per se*, but because of their health status combined with living alone, which for some individuals is a personal choice. As we develop models using richer individual-level data, these kinds of questions will need to be addressed if resources are to be allocated in a manner consistent with need. (See also (Schokkaert & van de Voorde 2003) for a discussion of related ideas).

Lastly, needs-based models must be validated. This is a huge challenge given the absence of a gold standard distribution of need against which to assess such models. Commonly employed methods of needs-adjustment can result in allocations that deviate from need more than do current allocations(Hutchison et al 2003). This is a problem that cannot be addressed through traditional model validation techniques such as split samples (though such techniques may play a role in developing valid utilization models that underlie needs-based funding formulae). Rather, it requires development of region-specific needs measures calculated independently of the needs-based model (perhaps using data not routinely available and which therefore are not suitable for use in formulae themselves), against which the results of the needs-based model can be assessed. Although these challenges are formidable, we believe that creative solutions can be found and that analyses like the one presented herein offer considerable potential to improve the allocation of public health care resources.

Table 1: Independent Variables: Description and Rationale for Inclusion

Variable	Description and Rationale	Need Status*
Demographics		
Age	Respondent's age at time of interview. Age is strongly associated with chronic health decline and need for health care.	N
Sex	Sex of respondent. Males and females have different lifetime profiles of health care need and consumption.	N
Language	Indicator of whether respondent can speak English. Those who cannot speak English may experience barriers to care.	NN
Aboriginal	Indicator of whether respondent is Aboriginal. Aboriginals are marginalized members of Canadian society, who on average have both higher needs and reduced access to health care services	C
Health/Functional Status		
Self-assessed health status	Respondent's self-assessed health status (Excellent, Very Good, Good, Fair, Poor). SAHS is a well-validated general health measure.	N
Hospital Admissions	Respondent's number of hospital admissions during the study period. Hospital admissions reflect need for acute, post-hospitalization home care.	N
Chronic Conditions	Respondent's self-reported number of chronic conditions that are associated with need for home care. Chronic conditions are a significant risk factor for long-term need for home care services.	N
Activities of Daily Living	Number of up to five activities of daily living for which respondent requires assistance. Need for assistance with activities of daily living is a well-established risk factor for home care need.	N
GP/FP Visits	Respondent's number of general practitioner visits during the study period. Included as a residual health status measure; may also reflect respondent's general propensity to use health care services and MD-facilitated access to home care.	N
Informal Support		
Marital Status	Indicator of whether respondent is unmarried or married. Spouse can provide informal support, reducing need for home care.	N
Living Arrangement	Indicator of whether respondent lives alone or live with at least one other individual. Housemate(s) can provide informal support.	N

Table 1: Independent Variables: Description and Rationale for Inclusion

Contact with Neighbours	Indicator of degree of respondent's contact with neighbours (none, some, daily). Indicator of informal support available.	N
Social Support	An index of the strength of respondent's social support network. Health care need is, in general, related to social support.	N
Socio-economic Status		
Education	Highest level of education attained. Represents three potential influences: (1) general knowledge and ability to advocate and navigate system to obtain services; (2) unmeasured aspects of health status; (3) attitudinal influences associated with socio-economic status	C
Income	Household income per capita. Represents three potential influences: (1) ability to purchase private home care services; (2) unmeasured aspects of health status; (3) attitudinal influences associated with socio-economic status.	N
Aggregate Demand and Supply Factors		
Historical Budget	CCAC region's budget, adjusted for the age-sex distribution of the population in the CCAC region. Indicator of historical adequacy of CCAC funding, and in particular, the extent to which utilization in CCAC region may be supply constrained.	NN
Chronic Beds	Supply of chronic care hospital beds per capita in respondent's CCAC region. A potential substitute for home care services.	C
Nursing Home Beds	Supply of nursing home beds per capita in respondent's CCAC region. A potential substitute for home care services.	C
GP/FP supply	Supply of general/family practitioners per capita in CCAC region. Indicator of access to GP/FP services, which can play role in gaining access to home care services; may also correlate with general supply of health care resources in CCAC region	C
Acute-Care-Induced Home Care	The expected number of acute-care-induced days of home care utilization per capita in the CCAC, derived from CCAC-specific rates of hospitalization (by diagnostic code) and provincial-level association between acute care discharge (by diagnostic code) and home care use. Indicator of aggregate demand in a region emanating from activity in the acute-care hospital sector.	NN
Region Dummies	CCAC-specific regional dummy variables to capture unobservable region-specific effects.	NN

*N = need-related; NN = non-need-related; C = classification is contingent on the empirical relationship identified

Table 2: Descriptive Statistics

Variable	Mean	s.d.	Range	Variable	Mean	s.d.	Range
Dependent Variables							
Use vs non-use	0.06	0.25	0-1				
Expenditure per user	3755.0	6334.00	0- 52430				
Independent Variables							
Age	44.80	19.00	12-102	Social Support Index			
Female	0.53	0.50	0-1	Level 1 (Least support)	0.01	0.10	0-1
Self-Assessed Health				Level 2	0.02	0.12	0-1
Excellent Health	0.24	0.43	0-1	Level 3	0.03	0.16	0-1
Very Good Health	0.39	0.49	0-1	Level 4	0.07	0.25	0-1
Good Health	0.25	0.44	0-1	Level 5 (Most support)	0.88	0.32	0-1
Fair Health	0.09	0.28	0-1	Educational Attainment			
Poor Health	0.03	0.17	0-1	< Grade 9	0.11	0.31	0-1
Chronic Conditions				Some Secondary	0.22	0.41	0-1
< 4 Chronic Conditions	0.98	0.37	0-1	Secondary Grad	0.68	0.47	0-1
4+ Chronic Conditions	0.02	0.01	0-1	Household Income Per Cap			
Activities of Daily Living (ADLs)				Income Quartile1	0.24	0.43	0-12K
Requiring Assistance				Income Quartile 2	0.23	0.42	12K-16K
ADL 0	0.93	0.25	0-1	Income Quartile 3	0.24	0.43	16K - 24K
ADL 1	0.03	0.16	0-1	Income Quartile 4	0.29	0.46	24K+
ADL 2	0.01	0.12	0-1	Aboriginal Status	0.01	0.09	0-1
ADL 3	0.01	0.11	0-1	English Language	0.99	0.09	0-1
ADL 4	0.01	0.08	0-1	Chron Hosp Beds/10000	8.80	3.99	2.0-19.0
ADL 5	0.01	0.09	0-1	LT Care Beds/10000	55.55	18.26	19 - 98
GP Visits	15.33	17.76	0 - 273	Acute-Care-Induced Home-			
Hospital Admissions				Care Days per Capita	0.24	0.05	0.15-0.37
0 Hospital Admissions	0.69	0.46	0-1	Adjusted Budget	1.02	0.12	.91 - 1.44
1 Hospital Admission	0.17	0.38	0-1				
2+ Hospital Admissions	0.13	0.34	0-1				
Marital Status (Unmarried = 1)	0.46	0.50	0-1				
Living Arrangement (Alone = 1)	0.22	0.41	0-1				
Contact with Neighbours							
None	0.14	0.35	0-1				
Some	0.61	0.49	0-1				
Daily	0.25	0.43	0-1				

Table 3: Part 1 Logistic Regression Results (N = 22855)

Dependent Variable: 0 = No Home Care; 1 = Home care User

	Model: Age-Sex	Model: Full
pseudo-R^2	0.19	0.41
Difference in predicted between users and non-users:	0.15	0.34
Log-likelihood:	-3999.4	-2499.5

Variable	Odds Ratio	p	Variable	Odds Ratio	p
Age	0.950	0.00	Age_ADL3	1.028	0.02
Age2	1.001	0.00	Age_ADL4	1.022	0.05
Female	0.530	0.04	Age_ADL5	1.007	0.53
age_fem	1.016	0.00	Age_GPvisits	1.000	0.00
Vg hlth	1.393	0.03	1 Hosp Adm	17.267	0.00
Good hlth	1.366	0.02	2+ Hosp Adm	116.081	0.00
Fair hlth	1.827	0.00	Age_hosp1	0.982	0.01
Poor hlth	3.104	0.00	Age_hosp2	0.971	0.00
4+ chronic conditions	0.009	0.01	Sex_hosp1	0.901	0.68
Age_chronic 4+	1.063	0.01	Sex_hosp2	0.600	0.02
1 ADL	0.503	0.39	Age_alone	1.008	0.00
2 ADL	3.575	0.08	Educ < grade 9	2.222	0.01
3 ADL	0.587	0.53	Age_ed< g9	0.989	0.02
4 ADL	1.798	0.43	Inc - Quart 1	0.386	0.00
5 ADL	7.068	0.09	Inc - Quart 2	0.601	0.20
Age_ADL1	1.019	0.08	Age_inc1	1.020	0.00
Age_ADL2	1.000	0.84	Age_inc2	1.012	0.07

Factor	Change in Prob of Using Home Care*: Age 30	Age 60	Age 90
Sex: female vs. male	0.000	0.004	0.082
SAHS: very good vs. excellent	0.001	0.001	0.001
SAHS: good vs. excellent	0.001	0.002	0.008
SAHS: fair vs. excellent	0.002	0.004	0.020
SAHS: poor vs. excellent	0.007	0.010	0.050
ADLs: 1 ADL vs. 0 ADLs	-0.001	0.004	0.104
ADLs: 2 ADL vs. 0 ADLs	0.008	0.020	0.139
ADLs: 3 ADL vs. 0 ADLs	0.000	0.015	0.281
ADLs: 4 ADL vs. 0 ADLs	0.010	0.050	0.442
ADLs: 5 ADL vs. 0 ADLs	0.026	0.079	0.460
Hospitalizations: 1 hosp adm vs. 0 hosp adm	0.032	0.045	0.198
Hospitalizations: 2+ hosp adm vs. 0 hosp adm	0.126	0.127	0.312
Living Arrangement: living alone vs. living with someone	0.001	0.007	0.081
Education Level: Less than grade 9 vs. grade 9 or higher	-0.002	-0.001	0.013
Household Income Level: 1st quartile vs. above median	-0.001	0.002	0.047
Household Income Level: 2nd quartile vs. above median	-0.001	0.001	0.027

*Base case: male, unmarried, excellent health, living alone, 0 chronic conditions, 0 ADLs, no hosp, secondary grad, 3rd quartile income

Table 4: Part 2 GLM Model of Home Care Expenditures

Conditional on Use (N = 1447)

Dep Var: $ Value of Home Care Services Received		
	Age-Sex	**Full Model**
Scaled Deviance	2.16	1.08
Akaike Information Criterion	18.3	12.3

Variable	**Coeff**	**p**
Female	-1.072	0.00
Age_fem	0.016	0.00
Single	0.543	0.00
4+ chron	-2.493	0.09
Age_chr4+	0.034	0.06
1,2,3 ADL	0.816	0.00
4 ADL	1.296	0.00
5 ADL	1.985	0.00
Educ < G9	-0.250	0.01
Age_GPvst	0.000	0.00
1 Hosp	-0.690	0.05
2+ Hosp	0.881	0.01
Age_hosp1	0.011	0.04
Age_hosp2	-0.009	0.03

	Difference in Expenditure	
	Age = 30	**Age = 90**
Factor		
Sex: Female vs. male	-714	809
Marital Status: Single vs. Married	1020	1103
Chronic Condition: < 4 vs. 4+	-1238	1484
ADL: 1,2,3 vs. 0	2031	2195
ADL: 4 vs. 0	4250	4594
ADL: 5 vs. 0	10094	10913
Education Level: < G9 vs. G9 +	-580	-627
Hospitalization: 1 vs. 0	-482	500
Hospitalizations: 2+ vs. 0	1318	44

Table 5: Need-Adjusted, Risk-Adjusted and Actual Capitation Rates (1999-00 Cdn $)

CCAC Region	Need-Adj Cap Rate	Need-Adj 95% CI	Actual Cap Rate	% Diff Act vs. Need	Risk-Adj Cap Rate	Risk_Adj 95% CI	% Diff Risk vs. Need
12*	55	[35, 67]	111	104.0%	41	[8,94]	-25.5%
30	61	[35,81]	75	23.4%	63	[27,100]	4.3%
5*	66	[38,95]	124	88.3%	41	[13,68]	-37.6%
16	68	[39,96]	69	1.4%	53	[27,81]	-22.5%
17*	74	[51,98]	99	34.3%	57	[31,87]	-22.2%
7	77	[45,108]	91	18.6%	51	[12,100]	-34.1%
39*	83	[55,122]	148	78.7%	126	[65,228]	51.9%
35	84	[48,121]	114	35.4%	101	[54,169]	19.4%
22	85	[59,112]	98	15.5%	78	[41,126]	-8.5%
21	89	[56,123]	93	4.9%	79	[35,135]	-10.9%
13	90	[40,153]	112	23.8%	68	[23,133]	-24.7%
38*	91	[62,124]	126	38.4%	124	[74,194]	36.5%
20	93	[54,153]	102	9.7%	83	[41,141]	-11.3%
15	99	[63,146]	102	2.8%	76	[43,130]	-22.9%
36	102	[74,138]	122	19.3%	122	[62,210]	19.4%
25	104	[70,146]	125	20.7%	97	[54,160]	-6.1%
33	105	[55,178]	167	58.8%	119	[48,230]	13.6%
31	113	[74,160]	111	-1.2%	126	[59,211]	11.7%
11	114	[75,164]	123	7.9%	85	[43,155]	-25.7%
34	115	[86,157]	104	-9.8%	133	[88,205]	16.2%
40	118	[78,166]	128	8.6%	188	[111,307]	59.6%
24	120	[69,189]	134	11.7%	112	[47,210]	-6.6%
27	123	[91,166]	137	10.8%	124	[70,193]	0.1%
42	126	[62,203]	143	14.3%	210	[95,346]	67.1%
43	126	[71,200]	157	24.9%	250	[64,612]	98.5%
37	127	[70,197]	129	1.4%	162	[62,340]	27.5%
23	138	[93,199]	112	-19.2%	129	[72,211]	-6.9%
6	144	[100,215]	129	-10.5%	94	[56,155]	-34.7%
29	144	[90,230]	146	1.3%	148	[79,261]	3.1%
14	146	[84,230]	152	4.2%	113	[28,249]	-22.9%
41	147	[90,232]	143	-2.6%	235	[78,474]	59.8%
19	151	[85,248]	124	-18.0%	125	[61,227]	-17.6%
1	152	[72,271]	148	-3.1%	64	[1,169]	-58.0%
2	161	[49,339]	117	-27.3%	82	[22,152]	-49.3%
18	176	[119,258]	141	-19.8%	144	[73,261]	-17.7%
32	184	[110,291]	142	-22.8%	208	[100,368]	13.4%
9	202	[44,532]	132	-34.8%	141	[29,298]	-30.1%
26	233	[86,457]	167	-28.4%	228	[82,466]	-2.3%
8	241	[126,429]	138	-42.5%	162	[66,333]	-32.7%
10	241	[124,422]	134	-44.6%	173	[87,322]	-28.1%
28	275	[56,722]	100	-63.6%	276	[54,759]	0.2%
3	318	[89,687]	150	-52.9%	179	[34,427]	-43.7%
4*	347	[130,667]	123	-64.5%	206	[73,408]	-40.8%

* CCAC region's actual cap rate falls outside the needs-based funding 95% confidence interval
All capitation rates based on 1999-00 budget

Table 6: Optimal Allocation of 5% Budget Increase to CCAC Regions

CCAC Region	Actual Cap Rate (1)	Need-Adj Cap Rate (2)	Derivative Statistic (1/2)	95% CI	Cap Rate after Budget Increase	Change
4	123	347	0.355	[0.19,0.95]	285	162
28	100	275	0.364	[0.14,1.78]	226	126
3	150	318	0.471	[0.22,1.67]	261	111
10	134	241	0.554	[0.32,1.07]	198	64
8	138	241	0.575	[0.32,1.10]	198	60
9	132	202	0.652	[0.25,2.97]	166	34
26	167	233	0.716	[0.37,1.88]	191	24
2	117	161	0.727	[0.34,2.39]	132	15
32	142	184	0.772	[0.49,1.29]	151	9
18	141	176	0.802	[0.54,1.18]	144	3
23	112	138	0.808	[0.56,1.21]	113	1
19	124	151	0.820	[0.50,1.46]	124	0
6	129	144	0.895	[0.60,1.28]	129	0
34	104	115	0.902	[0.66,1.20]	104	0
1	148	152	0.969	[0.54,2.05]	148	0
41	143	147	0.974	[0.61,1.58]	143	0
31	111	113	0.988	[0.70,1.50]	111	0
29	146	144	1.013	[0.63,1.62]	146	0
37	129	127	1.014	[0.66,1.84]	129	0
16	69	68	1.014	[0.72,1.76]	69	0
15	102	99	1.029	[0.70,1.61]	102	0
14	152	146	1.042	[0.66,1.82]	152	0
21	93	89	1.049	[0.76,1.65]	93	0
11	123	114	1.079	[0.75,1.64]	123	0
40	128	118	1.086	[0.77,1.64]	128	0
20	102	93	1.097	[0.67,1.88]	102	0
27	137	123	1.108	[0.82,1.51]	137	0
24	134	120	1.118	[0.70,1.95]	134	0
42	143	126	1.143	[0.71,2.31]	143	0
22	98	85	1.155	[0.88,1.66]	98	0
7	91	77	1.187	[0.85,2.04]	91	0
36	122	102	1.194	[0.89,1.65]	122	0
25	125	104	1.207	[0.86,1.80]	125	0
30	75	61	1.234	[0.92,2.17]	75	0
13	112	90	1.238	[0.73,2.79]	112	0
43	157	126	1.249	[0.79,2.22]	157	0
17	99	74	1.343	[1.00,1.95]	99	0
35	114	84	1.354	[0.94,2.37]	114	0
38	126	91	1.384	[1.01,2.02]	126	0
33	167	105	1.588	[0.94,3.04]	167	0
39	148	83	1.787	[1.21,2.69]	148	0
5	124	66	1.883	[1.30,3.27]	124	0
12	111	55	2.040	[1.66,3.19]	111	0

Appendix 1

Hurley (2003) provides full details. Assume the following: a society of R individuals, each of whom lives in a different region, with identical preferences defined over health status and other goods, but differing needs for health care (where need is defined as the amount of health care required to exhaust benefit). The specification of utility functions and the social welfare function which leads to the optimality of the relative-need allocation principle is as follows:

Government Problem:

$$\underset{HC_1 \ldots HC_R}{Max} - \sum_{r=1}^{R} \left| \frac{N_r}{N} \left(\frac{HC_r}{N_r} - 1 \right)^{\delta} \right| \qquad s.t. \quad \sum_{r=1}^{R} HC_r \leq Y \qquad \text{A1.1}$$

Where

N_r	= level of need for individual r
N	= total society need
HC_r	= level of health care funding in region r
Y	= health care budget
δ	= parameter (δ ≥ 1) indicating the steepness and concavity of the relationship between deviations from need and utility

The solution to this optimization problem is:

$$HC_r^* = \frac{N_r}{N}(Y) \qquad \text{A1.2}$$

A1.2 is the relative-need allocation rule.

To allocate incremental funds optimally, calculate the derivative of the social welfare function, evaluated at each region's current level of funding (HC_r). The region with the largest derivative has the highest priority for additional funding. Letting W denote the social welfare function from A1.1 above, the relevant derivative is as follows:

$$W_{Y_r} \quad = -\frac{\delta}{N}\left(\frac{HC_r}{N_r} - 1 \right) \qquad \text{A1.3}$$

Unfortunately, we do not observe δ, N or N_r . $\frac{\delta}{N}$, however, is just a constant of proportionality that factors out as it does not vary across regions. Furthermore, the ratio $\frac{HC_r}{N_r}$ can be estimated. The current budget, Y, is some proportion of total need (Y = αN, α ≤ 1)). Based on this fact, and assuming that the empirical model is valid, it follows that the estimated needs-based capitation rates (HC_r^*) are the same proportion, α, of each region's need. That is, HC_r^* is an estimate of αN_r. This can be substituted into A1.3 to obtain:

$$\frac{1}{\alpha}\hat{W}_{Y_r} = -\frac{\delta}{N}\left(\frac{HC_r}{HC_r^*} - 1\right) \quad \text{A1.4}$$

From this if follows that the funding should go first to that region for which

$$\frac{HC_r}{HC_r^*} < \frac{HC_s}{HC_s^*} \text{ for all } s \neq r \quad \text{A1.5}$$

We call $\frac{HC_r}{HC_r^*}$ the "derivative statistic". We can construct 95% confidence intervals around our point estimate based on the standard error of HC_r^*:

$$se\left(\frac{HC_r}{HC_r^*}\right) = HC_r se\left(\frac{1}{HC_r^*}\right) \quad \text{A1.6}$$

References

Blough K, Madden C, Hornbrook M. 1999. Modeling Risk using Generalized Linear Models. *Journal of Health Economics* 18(2):153-71

Culyer A, van Doorslaer E, Wagstaff A. 1992a. Access, Utilization and Equity: A Further Comment. *Journal of Health Economics* 11(2):207-10

Culyer A, van Doorslaer E, Wagstaff A. 1992b. Utilization as a Measure of Equity by Mooney, Hall, Donaldson and Gerard: Comment. *Journal of Health Economics* 11(1):93-8

Duan N. 1983. Smearing Estimate: A Non-Parametric Retransformation Method. *Journal of the American Statistical Association* 78:605-10

Hauck K, Shaw R, Smith P. 2001. Reducing Avoidable Inequalities in Health: A New Criterion for Setting Capitation Payments. *Health Economics* 11:667-77

Hurley J. 2003. An Economic Model of Needs-based Funding in Health Care. *Unpublished*

Hutchison B, Torrance-Rynard V, Hurley J, Birch S, Eyles J, Walter S. 2003. Equity in Health Care Funding: Comparison of Expenditures in Ontario Canada to Allocations based on Population Need. Unpublished

Jones A. 2000. Health Econometrics. In *Handbook of Health Economics*, ed. Culyer A, Newhouse J, Amsterdam: Elsevier.

Kahneman D, Varey C. 1991. Notes on the Psychology of Utility. In *Interpersonal Comparisons of Well-being*, ed. Elster J, and Roemer J. Cambridge: Cambridge University Press. 127-163 pp.

Kennedy P. 1986. Interpreting Dummy Variables. *Review of Economics and Statistics* 68(1):174-5

Manning W. 1998. The Logged Dependent Variable, Heteroskedasticity, and the Retransformation Problem. *Journal of Health Economics* 17(3):283-95

Manning W, Mullahy J. 2001. Estimating Log Models: To Transform or Not to Transform? *Journal of Health Economics* 20(4):461-94

Mooney G, Hall J, Donaldson C, and Gerard K. 1992. Reweighing Heat: Response to Culyer, van Doorslaer and Wagstaff. *Journal of Health Economics* 11(2):199-206

Mooney G, Hall J, Donaldson C, Gerard K. 1991. Utilization as a Measure of Equity: Weighing Heat? *Journal of Health Economics* 10(4):383-476

Mullahy J. 1998. Much Ado about Two: Reconsidering Retransformation and the Two-Part Model in Health Econometrics. *Journal of Health Economics* 17(3):247-81

Rice N, Smith P. 2001. Capitation and Risk Adjustment in Health Care Financing: An International Progress Report. *Milbank Quarterly* 79(1):81-113

Schokkaert E, van de Voorde C. 2003. Risk Selection and the Specification of the Conventional Risk Adjustment Formula. *Journal of Health Economics* Forthcoming

Smith P, Rice N, Carr-Hill R. 2001. Capitation Funding in the Public Sector. *Journal of the Royal Statistical Society, A* 164(2):217-57

Statistics Canada. 2002. *Variance Estimation Using Bootstrap Weights: Programmers Guide*. Ottawa: Statistics Canada.

Suits D. 1984. Dummy Variables: Mechanics vs Interpretation. *Review of Economics and Statistics* 66(1):177-80

van de Ven W, Ellis R. 2000. Risk Adjustment in Competitive Health Care Plans. In *Handbook of Health Economics*, ed. Culyer A, and Newhouse J, Amsterdam: Elsevier.

Wagstaff A, and van Doorslaer E. 1993. *Equity in the Finance and Delivery of Health Care: An International Perspective*. Oxford: Oxford University Press.

Williams A. 1978. Need - An Economic Exegesis. In *Economic Aspects of Health Services*, ed. Culyer A, Wright K. London: Martin Robertson. 32-45 pp.

Yaari M, Bar-Hillel M. 1984. On Dividing Justly. *Social Choice and Welfare* 1(1):1-24