Assessing the Regularity and Predictability of the Age-Trajectories of Healthcare Utilization

by

Margaret Turnbull

Submitted in partial fulfilment of the requirements
for the degree of Master of Science

at

Dalhousie University
Halifax, Nova Scotia
August 2012

DALHOUSIE UNIVERSITY

DEPARTMENT OF COMMUNITY HEALTH AND EPIDEMIOLOGY

The undersigned hereby certify that they have read and recommend to the Faculty of

Graduate Studies for acceptance a thesis entitled "Assessing the Regularity and

Predictability of the Age-Trajectories of Healthcare Utilization" by Margaret Turnbull in

partial fulfilment of the requirements for the degree of Master of Science.

Dated: August 20th 2012

Supervisor: _____

Readers: _____

_____

DALHOUSIE UNIVERSITY

DATE: August 20, 2012

AUTHOR: Margaret Turnbull

TITLE: Assessing the Regularity and Predictability of the Age-Trajectories of Healthcare Utilization

DEPARTMENT OR SCHOOL: Department of Community Health and Epidemiology

DEGREE: MSc   CONVOCATION: October   YEAR: 2012

_____
Signature of Author

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

This research examines the viability of a need-based approach that models the age-trajectories of healthcare utilization. We propose a fundamentally different way of treating age in modeling healthcare use. Rather than treating age as a need indicator, we refocus modeling efforts to predicting the age-trajectories of healthcare use. Using inpatient hospital utilization data from the Discharge Abstract Database, first, we model the age-trajectories of the rate of hospital use employing a common functional form. Second, we assess variation in these age-trajectories using growth curve modeling. Third, we explain variation in these age-trajectories using census variables. Our analysis shows that the regional variation in the age-trajectories of the rate of inpatient hospital use is sufficient to justify this method, and could be partially explained using census variables. This indicates that modeling age-trajectories of healthcare use is advantageous, and the current need-based approach may benefit from this new modeling strategy.

LIST OF ABBREVIATIONS USED

CD   Census Division

CIHI   Canadian Institute of Health Information

DAD   Discharge Abstract Database

DLI   Data Liberation Initiative

LOS   Length of Stay

SES   Socioeconomic status

ACKNOWLEDGEMENTS

CHAPTER 1: INTRODUCTION

Identifying how best to allocate healthcare resources is a central and continuous challenge for health policy makers. As healthcare funding has typically been distributed to regions based on their prior use, a region's ability to provide healthcare has been influenced by previous consumption, existing infrastructure, and politics, and not on the population's relative need. It is now increasingly accepted that using historical information does not lead to equitable allocation because it disregards the population's needs, may perpetuate existing inequalities, and may even encourage unnecessary use. Demands for more equitable healthcare systems, particularly in publicly funded healthcare systems, have led governments' interest in approaches that allocate healthcare funding based on measures of relative need. These models are commonly referred to as need-based resource allocation models.

Considerable research and political efforts have motivated the development of need-based resource allocation models for healthcare planning, and most of these need-based approaches employ a common analytical framework to define and adjust for need (1). The common analytical framework is characterized by the following basic methodology: standard levels of healthcare use are determined, and then applied to population data to estimate expected utilization based on population need. To develop a standard level of healthcare use, actual healthcare utilization is first modeled as a function of need and non-need factors, and then models are used to predict need-expected use while purging the influence of non-need factors (2). Although this analytical framework is commonly used and provides many advantages, it suffers from a number of limitations that have rendered it impractical in many jurisdictions. One such limitation is the treatment of age within the models.

Age has been used as a fundamental indicator of need in virtually all need-based approaches. Healthcare need is strongly correlated with age, as age is a proxy for both morbidity and mortality. However, age is not a direct cause of morbidity or mortality. Instead, its effects on morbidity and mortality reflect the accumulation of health deficits

associated with life course experiences and exposures that lead to the development of chronic disease and acute illness. The age pattern of morbidity and mortality in a population reflects the effects of health indicators on birth cohorts as they age. As a result of regional differences in determinants of health, the age pattern of morbidity and mortality is unlikely uniform across regions. Lack of uniformity should be reflected in differences in the age-trajectories of need for healthcare services. As has been shown with decades of research examining variations in the age-trajectories of mortality across populations (3–5), we expect that the age-trajectories of healthcare use are characterized by variable but predictable patterns in healthcare use by age, particularly once adjusted for the effects of non-need indicators of health services use.

Accordingly, we proposed that it is not appropriate to simply include age as a need indicator when developing need-based approaches. Rather, modeling efforts should be refocused on understanding and predicting the age-trajectories of healthcare use as a function of need and non-need determinants. The first step in developing new need-based approaches should refocus on modeling the age-trajectories of healthcare use by age, and not on simply modeling healthcare use with age as one of a series of need indicators. We evaluated whether the trajectories of healthcare use by age followed predictable patterns that varied between regions, as a function of differences in population health needs. As a first step in creating such a model, this project assessed whether:

A. the regional age-trajectories of inpatient hospital use could be modeled with a common functional form;

B. there was significant regional variation in the age-trajectories of inpatient hospital use; and

C. the regional variation in the age-trajectories of inpatient hospital use could be partially explained by ecological-level need indicators

To assess these questions, we used inpatient hospital utilization data from the Discharge Abstract Database for Canadian census divisions, and proceeded with the following three steps. First, we graphed and described the relationship between hospital utilization and

age, by census division, to identify a functional form that was appropriate for modeling age-trajectories of the rate of inpatient hospital use. Second, we used growth curve modeling to model the regional variation in the age-trajectories of the rate of inpatient hospital use. Third, we assessed the degree to which regional variations in the age-trajectories of the rate of inpatient hospital use could be predicted using ecological-level measures of need from the 2006 Census.

The thesis is organized into five chapters. In chapter two, the background, we discuss the conceptual justification for proposing this methodology. Chapter three outlines the objectives of this paper. Chapter four consists of the manuscript, which is a general overview of the methodology, results, and discussion. Chapter five provides additional discussion and conclusion, which further details the implications and relevance of the findings.

CHAPTER 2: BACKGROUND

Over the past 50 years, there has been much discussion worldwide around what is the most efficient and fair way to allocate limited healthcare resources. As many countries experience rising healthcare expenditures, the pressure mounts to identify an acceptable way to distribute valuable healthcare dollars. Many developed countries seek to promote equitable allocation of healthcare resources with need-based resource allocation models (6,7). While a great deal of progress has been made in the last 50 years in both health services and medicine worldwide, a need-based resource allocation model that can be considered as the gold standard has not yet been developed. There remains a need to develop a need-based resource allocation model that can be implemented widely.

2.1 Equity

Equity is considered as one of the fundamental pillars of health policy in many countries, in spite of the difficulty in clearly defining and measuring it (6,8,9). In health service research, we typically define equity in the context of service access, where pursuing equity in access to health services reflects efforts to reduce unequal opportunities for health services access between persons or populations after taking into account their need for health services. Equitable access thus means equal access for equal need for health services regardless of the geographical, financial, educational, or cultural barriers. The idea of equal access for equal need also means unequal healthcare use for different levels of need. We expect that someone who is chronically ill, with lung cancer, for example, is going to need more services than someone who is suffering from a common cold (10). Healthcare access is inequitable when someone has greater access to healthcare for reasons not related to healthcare need. For example, it would be inequitable if one person received more healthcare services than another, on the basis of his/her income, despite the identical healthcare need between them. By extension, we would consider it inequitable if a high-income region received more healthcare resources than a low-income region with identical healthcare need.

2.2 Funding Models

2.2.1 Historically-Based Funding

The common approach to allocating healthcare resources in many jurisdictions is to distribute funds according to the previous year's expenditures, typically referred to as historical-based funding (11). This model allows the funder to make slight funding modifications to reflect economic trends (12). Historically-based funding models are advantageous insofar as there are few data requirements and they are conceptually simple. Rice and Smith observe that repeating patterns of allocation is practical on a short-term basis because it is an efficient use of existing infrastructure, but they go on to caution that service structures such as hospitals are often poorly distributed, and historically-based funding will perpetuate any inequitable patterns of resource distribution (13). Maynard and Ludbrook astutely simplified these models to "what you got last year, plus an allowance for growth, plus an allowance for scandals" (14). Simply put, historical-based funding models may be practical and simple in the short-term, but perpetuate pre-existing inequity in the long run. In spite of the long-standing concerns that these models perpetuate inequity in allocating healthcare resources, many provincial and regional levels of government in developed countries still allocate healthcare resources according to the previous year's budget (15).

2.2.2 Need-Based Funding

There has been increasing pressure to redistribute healthcare funds according to a population's need for healthcare (16–18). The United Kingdom (UK) was arguably the first country that made a concerted effort to move beyond historically-based funding models and use population-level need indicators in the late 1970s. Its funding formula was created by the Resource Allocation Working Party (RAWP) and is still known by this name. Since its inception, most industrialized countries have developed their own variation of a need-based allocation model to distribute healthcare funds. While there is some variability in methodology, most need-based approaches follow the five steps described in detail below.

2.3 Common Analytical Framework in Need-Based Resource Allocation Models: an Overview

Step 1

The first step is to describe the variation in healthcare use by need and non-need variables in the following form:

$$y_i = \alpha A_i + \beta X_i + \gamma Z_i + \varepsilon_i \quad (1)$$

Where y_i is the healthcare use of individual i; α is a vector of age coefficients; A_i is a vector of dummy variables for age; β and X_i are, respectively, coefficients and variables for need indicators; γ and Z_i are, respectively, coefficients and variables for non-need indicators; and ε_i is an error term. Ideally, equation (1) is estimated with individual level data. However, as will be addressed below, individual-level data are often not available, and ecological data are used instead.

Broadly, this equation says that healthcare use (y_i) is determined by a person's age (A_i), their healthcare need (X_i), and other indicators known to influence use (Z_i).

2.3.1 Age And Sex As Need Indicators

Age is almost universally included as a need indicator (13), although it is not a direct measure of need. Age is strongly associated with mortality and morbidity, and as such, is a good proxy for a broad range of needs. As people age they tend to require higher levels of preventative care, curative care, and palliative care services. The incidence and prevalence of most chronic conditions increases with age, and the risk of death increases with age. Moreover, age also captures shared life experiences between individuals in the same cohort. The association between age and healthcare need is not only biologically supported but is influenced by shared generational exposures (1,19).

For the reasons listed above, age is heavily relied upon as an indicator of need in need-based resource allocation models. Age coefficients always have statistically significant, large effect sizes regardless of other variables included in equation (1). Sex is similar to age in that it is also a proxy for need and is widely included in most models (20). Moreover, age and sex often interact as proxies for need, because, for example, the age-trajectories of morbidity are different in men and women (21).

2.3.2 Other Need Indicators

Need indicators are 'legitimate' determinants of need for healthcare use. An example of a need indictor is a measure related to an individual's health or disease status. We expect a patient with a chronic condition, such as heart disease, to require more health services than someone with a non-life-threatening acute condition, such as an ear infection. Ideally, we obtain a perfect measure of a person's health status, and use this as our need indicator. Such a perfect measure does not exist, and we use a variety of measures to reflect a person's need for healthcare. These can include upstream determinants of health such as socioeconomic status, measures of health risk behaviors (e.g., obesity, smoking, and physical inactivity), and measures of morbidity (e.g., self-assessment of overall health status, or prevalence of a chronic disease). When individual-level data are not available, ecological-level measures of need are often used.

2.3.3 Non-Need Indicators

Ideally, need for healthcare would fully explain healthcare use. However, other determinants not related to healthcare need also influence use. These are non-need indicators. Typical examples of non-need indicators include: a patient's access to care, their help seeking behaviors, and practice variation between physicians and healthcare institutions. Much like need, it is often difficult to measure non-need. Depending on what other variables are included in a model, a determinant can reflect either need or non-need.

2.3.3.1 Example: Education

Education can act as a need indicator, if it reflects differences in health status resulting from the 'upstream' determinants of health. People with lower levels of education are more likely to suffer from a range of chronic and acute conditions when compared to people with higher levels of education (22). When need is not well measured in a model, education therefore serves as a proxy for need. On the other hand, if need is well measured, education should be considered as a non-need determinant. Studies have shown that after adjusting for healthcare need people with higher education receive more care than their less educated counterparts (23). This may be because education indicates a person's ability to navigate the system, for example. On the basis of equal care for equal need, healthcare use should not depend on non-need indicators. Thus, healthcare use associated with education after adjustment for need is inequitable.

To summarize, the first step of the common analytical framework is to describe actual healthcare use. Actual utilization is influenced by need and non-need variables, but, as we saw with the example of education, characterizing variables as need or non-need is not always straightforward. Age is a unique variable because it is such a strong indicator of healthcare need.

Step 2

The second step of the common analytical approach is to estimate each individual's need-expected utilization ($\hat{y}_i^*$):

$$\hat{y}_i^* = \hat{\alpha}A_i + \hat{\beta}X_i + \hat{\gamma}\overline{Z}_i \quad (2)$$

As mentioned above, healthcare use reflects need as well as non-need indicators of utilization. To build a need-based resource allocation model, we would be interested in how much healthcare a person is expected to use based solely on his or her need. To estimate need-expected use, we purge the effects of non-need indicators by holding the values constant (often by setting them at their means) so that need indicators alone influence healthcare use.

Step 3

The third step is to estimate each region's per-capita healthcare need ($\hat{y}_r^{need}$). We do so by aggregating each individual's expected need-based utilization within each region and dividing it by its population, using sample weights.

$$\hat{y}_r^{need} = \frac{\sum \hat{y}_r^{need} \times W_{ri}}{\sum W_{ri}} \quad (3)$$

where w_{ri} is the weight assigned to individual i in region r.

Step 4

After estimating regional per-capita healthcare need, we estimate each region's relative need ($\tilde{y}_r^{need}$) as compared to other regions, by dividing each region's aggregate healthcare need by the total provincial need.

$$\hat{y}_r^{need} = \left(\frac{\hat{y}_r^{need}}{\sum \hat{y}_r^{need}} * 100 \right) - 100 \quad (4)$$

This equation yields the proportion of need a region has as compared to other regions.

Step 5

The fifth and final step is to calculate the relative regional budget by multiplying each region's relative need with the total healthcare budget.

$$y_r^{allocation} = \hat{y}_r^{need} * M \quad (5)$$

where M is the total healthcare budget to be allocated to regions.

2.4 Notes on the Common Analytical Framework

Equal per capita allocation used by some jurisdictions (for example, allocation of the Canada Health Transfer from the federal government to provinces and territories) is the simplest application of the common analytic framework described above. For equal per capita allocation, equations (1) and (2) do not incorporate any variables. Implicit assumption for this procedure is that need is equal across individuals. Thus, with equations (4) and (5), the healthcare budget is allocated based on population size alone.

Reliance on utilization data, a key feature of the common analytical framework, has been criticized in the literature. Estimating need-expected utilization assumes that it is possible to separate healthcare need from use, but distinguishing need indicators from non-need indicators is not a simple task. Thus, critics argue for measuring need directly by measuring morbidity and/or mortality, such as the standardized mortality ratio (SMR) or the rate of a specific condition (e.g., cardiovascular disease). This approach, sometimes called epidemiological approach (24), however, is not free from challenges. The first, and most prohibitive, limitation to this epidemiological approach is that it is unclear how health status indicators translate into healthcare resource requirements. If a region has an SMR of 1.8, for example, should it receive 80% more resources than a region with an SMR of 1? There is no evidence to suggest that SMR is proportional to healthcare resources required. Similarly, health status indicators typically used in the population-level allocation models do not indicate what types of services, such as primary care, specialists, or hospital services, would best suit to meet the needs of a region (2). Moreover, measures of morbidity or mortality are prone to the same issues that need-expected utilization faces, including data limitations and the problem of how to adjust for non-need indicators. Despite the imperfection, therefore, the common analytical framework using utilization data is arguably the most suitable and realistic need-based approach currently available.

2.5 Limitations of the Common Analytical Framework

The common analytic framework suffers from at least four limitations. They are: measuring need and non-need is challenging; the models only predict small amounts of variation in healthcare use; models are becoming increasingly complex; and age is not used to its full potential. While all four will be discussed here in detail, this project focused directly on the application of the fourth limitation, the use of age within the model.

2.5.1 The Problem of Measuring Need And Non-Need

Implementing the common analytical framework is difficult in practice for two reasons: first, we do not have 'gold standard' measures of need and non-need to use, and second, the data required for the models are often unavailable at the individual level.

The accuracy and ultimate utility of a need-based model is dependent on what variables are included in equations (1) and (2) above. Despite a wealth of literature, identifying the best way to measure the health care needs of a population remains as a difficult task. There are a variety of possible approaches to measure need, but the gold standard remains elusive (2,10). While age and sex are typically included in most models, there is significant variation in other need and non-need indicators used. Because there is no gold standard, concerns remain regarding the comparability and validity of different models. Kephart and Asada found that the use of different need indicators in need-based modeling could yield different, and even conflicting, results (2).

The common analytical framework ideally calls for individual level data. However, data are often unavailable at the individual level, so these models use ecological (small area) data, where indicators are measured at the level of a geographic region (25).

Mortality data are commonly employed under the assumption that age- and sex-standardized mortality ratios are correlated with both acute and long-term morbidity (16,17,26,27). The primary limitation to using mortality data is that the relationship between mortality and morbidity is unclear, made worse the fact that mortality data do not include non-fatal morbidity (28). Furthermore, mortality data become increasingly irrelevant as a need indicator as populations in developed countries age and the prevalence of long-term chronic diseases increases. Rather than mortality, measures of morbidity, typically using self-rated health status, have been more widely used as a proxy for health need. Morbidity, unlike mortality, follows a continuum - patients can range from mildly to severely ill. Thus, morbidity data are more detailed and flexible than mortality. We can measure morbidity through self-report or using specific morbidity measures (i.e., incident or prevalence rates of specific conditions). Additionally, proximity to death is an important additional indicator of healthcare use, and can be used as a need indicator (29). Data availability remains an issue both for morbidity and mortality data, especially at the individual level.

An alternative to using morbidity and mortality variables is to use more distal social determinants of health, which are often widely available at the small area level. Some social and economic conditions influence a person's health status. Socioeconomic status (SES), for example, is inversely correlated with the onset of most diseases and life expectancy (30,31). Determinants including income, education, housing, ethnicity, and gender all influence a person's health outcome. Furthermore, social determinants operate spatially, and are generally manifested in regional age-trajectories of mortality and morbidity (3,27). Using ecological variables that represent social determinants of health avoids reliance on individual level data while overcoming the limitations of mortality and morbidity data.

2.5.2 Models Only Predict Small Amounts of Variation

Need-based approaches have been criticized for their lack of ability to predict variations in health across individuals (32). One report by Rice and Smith noted that even models

that included several demographic and social determinants (e.g., age, sex, race, marital status, education, and income) were rarely able to explain more than a few percentage points of variation in utilization (32). Our ability to predict future healthcare need at the individual level is limited, and at best we are only able to explain roughly 20 percent of the individual variation (33,34). The way to increase a model's ability to explain the variation is to introduce additional need and non-need determinants, and/or econometric methods. However, building such an extensive model is limited by data restrictions and makes the model methodologically complex, which is the next limitation discussed in subsection 2.5.3.

2.5.3 Models Are Becoming Increasingly Complex

The efficacy of a need-based model in predicting variation between regions depends on the accuracy of the indicators used in the model. One way to increase a model's predictability is to increase the number of predictive variables in the model, and employ econometric methods to adjust for unobserved variables, measurement error and endogeneity (35). Funding models based on the Resource Allocation Working Party often started with a population size count, the age distribution of the population, and a measure of morbidity. To increase the value of these models, more need and non-need indicators were included (36), along with more complex econometric estimation approaches of increasing complexity.

The problem here is twofold: first, the data requirement for these models is unruly. Often the data are only available at the national or sub-national (e.g., provincial or state) level, not at the small area or individual level, and cannot be extended to other jurisdictions. National governments have had more success with implementation because of the availability of national-level data; however, sub-national (e.g., provincial) governments have data restrictions that render these models futile at the regional level. Second, the more complex these models become, the less transparent they are. For a model to be implemented into practice it needs to be politically acceptable. It should be intuitive and relatively easy to understand and justify (37). Peacock and Segal noted that if an

allocation model was made publicly available and was sufficiently transparent, it would encourage democratic debate about its appropriateness, ultimately adding to its utility and worth (38).

2.5.4 Age Is Not Used to its Maximum Potential

Age is the most widely used indicator of healthcare need, and typically, unadorned need-based approaches rely exclusively on age as a need indicator. As discussed above, the general reliance on age as a need-indicator can be attributed to its strong association with both morbidity and mortality. Age is not a direct indicator of morbidity or mortality, but rather, it is an indirect proxy for healthcare need.

While age is strongly correlated with healthcare utilization, the use of age as a proxy assumes that the underlying relationship between healthcare need and morbidity or mortality across all regions is adequately described by the average relationship (39). It assumes that two populations with identical age distributions, after adjustment for other need and non-need variables in the models, should have the same healthcare need regardless of differences in morbidity and mortality between these populations.

However, assuming invariance of age effects on healthcare need is neither intuitive nor accurate. There is a wealth of evidence to suggest that different populations have predictably different age-trajectories of morbidity and mortality. For example, Canadian Aboriginal and American Native communities suffer disproportionately more from a number of chronic and acute illnesses, and tend to experience these diseases at an earlier age than their non-Aboriginal counterparts (40–42). In Canada cardiovascular disease is the leading cause of death within both Aboriginal and non-aboriginal Canadians (43). However, age-standardized mortality rates among Aboriginal women are 61% higher compared to non-Aboriginal women. Additionally, the stroke mortality rate is 44% and 93% higher among Aboriginal men and women (respectively) than the general Canadian population (44,45).

Consider two hypothetical populations of equal population size: population A with a high socioeconomic status, low age-standardized rates of morbidity, and high proportion of older individuals; and population B with a low socioeconomic status, high age-standardized rates of morbidity, and high proportion of younger individuals. Figure 1 illustrates the relationship between age and morbidity for these two hypothetical populations. A typical allocation model, using the common analytic framework explained above, and using only age as a need indicator, assumes the relationship between age and healthcare use would be uniform across populations. Such an allocation model would distribute more funds to population A, even though population B has a greater need for care.

The assumption that the relationship between age and healthcare need is constant across regions would, therefore, systematically bias against areas where morbidity and mortality are concentrated among younger ages (for example, population B in the example above). Regions with earlier onset of morbidity would have their need under-estimated and would thus be unable to provide the equal access to healthcare for equal need.

Even models that use need indicators in addition to age face the same problem. These models treat age as a fixed effect that could potentially be explained away should the model include a comprehensive range of need and non-need determinants. The problem is that in spite of considerable research efforts, researchers have not thought about age beyond its strong association with healthcare use. No matter how many other variables we include, age remains one of the most significant and influential indicators of healthcare service utilization. As long as age remains in the models as a fixed effect, models assume a uniform relationship between age and healthcare use across regions.

An alternative way of thinking about the subject is that health indicators influence the different age progressions of morbidity and mortality. That is, the age determined rate and frequency with which people will require healthcare follows a trajectory reflecting the historical and present influence of health determinants in the population. Viewed from this perspective, age should not be treated as a need indicator. Instead of modeling

utilization as the first stage in the standard approach, with age and sex as need indicators, we should focus on modeling the age-trajectories of utilization as a function of need indicators. We expect the age-trajectories of healthcare need and use to vary across populations, and have a strong geographical component.

2.6 Time for a Change

To summarize, the common analytical framework suffers from at least four limitations that have prohibited its use in health policy: that need and non-need are not directly measured; that the common analytical framework can only predict small amounts of variation; that the models are complicated and not transparent; and that age is not optimally used. In spite of continuing research efforts, the current models based on the common analytical framework have not overcome these limitations and, as a result, have not effectively been translated into practical use. It is possible that current research efforts are limited by the methodology, and a new approach would provide elucidation to the recurrent problems with the common analytical framework. As a way to overcome the noted limitations, we have explored a new approach to modeling healthcare utilization as a first stage in developing need-based resource allocation models.

While this project was motivated by the need for an improved need-based resource allocation model, the new methodology proposed here can extend to all need-based approaches. For that reason, this project highlights the implications of modeling the age-trajectories in all need-based approaches, and is not restricted to need-based resource allocation models.

We proposed that a need-based approach that models the age-trajectories of healthcare utilization better uses age in the model than the current need-based approaches that include age as a need-indicator. The age-trajectories of healthcare need, as reflected in healthcare use, should be modeled and conceptualized as the accumulated influence of health indicators operating in a population. By approaching the need-based modeling in this way, we conceptualized both the role of age and the mechanisms by which health

indicators determine healthcare need. This approach directly addressed the fourth limitation we have identified: how age has been handled; however, we also believe it has potential to indirectly address other limitations as well. By focusing on modeling the age-trajectories of healthcare need, rather than healthcare use at the individual level, the variation explained by models may be dramatically improved. This approach may also more clearly articulate the influence of health indicators in a way that is consistent with the life course perspective, which has come to dominate the literature in this area. Finally, the approach is intuitively appealing and this may lead to more transparent models.

Refocusing the modeling on the age-trajectories of need can be accomplished through the use of readily available statistical methods. Specifically, we employed growth curve modeling for this purpose. Growth curve models have now been extensively employed in the social and biological sciences to model aging and growth processes, and their determinants. For example, growth curve models have been used to describe and predict the growth of children, patterns of child development and aging (46), and tumor growth (47,48). These models are readily estimated using standard statistical software, while opening the door to an extensive literature with potential to enhance and improve their application to needs based modeling.

To summarize, this thesis explored the viability of a need-based approach that models the age-trajectories of healthcare utilization.

CHAPTER 3: OBJECTIVES

The aim of this project was to illustrate and evaluate the viability of a need-based approach that modeled the age-trajectories of healthcare utilization. We proposed a fundamentally different way to treat age in modeling healthcare. Rather than treating age as a need indicator, we have refocused modeling efforts to understanding and predicting the age-trajectories of healthcare use. To do this, the project had three objectives:

1) To evaluate whether the regional age-trajectories of inpatient hospital use could be modeled with a common functional form;
2) To evaluate whether there was significant regional variation in the age-trajectories of inpatient hospital use; and
3) To determine whether the regional variation in the age-trajectories of inpatient hospital use could be partially predicted by ecological-level need indicators

CHAPTER 4: MANUSCRIPT

A New Approach to Modeling Healthcare Utilization and Need as a Function of Age

4.1 Introduction

Need-based approaches, which estimate healthcare resource requirements based on the needs of a population, are among the central methods in health services research and policy analysis. Their application is widespread and includes healthcare resource allocation (1), health human resource planning (7,49,50), analysis of equity in healthcare use (51,52), hospital performance evaluation (53,54), and risk adjustment (55). Need-based approaches typically begin by modeling healthcare utilization as a function of need indicators – 'legitimate' determinants of healthcare use –, while purging the influence of non-need indicators – 'illegitimate' determinants of healthcare use (10). In all applications, healthcare utilization is strongly predicted by age; so much so that unadorned need-based approaches often rely exclusively on age as a predictive variable. The general reliance on age as an indicator of healthcare utilization can be attributed to its strong association with both morbidity and mortality. In addition, the association between age and healthcare need is not only biologically supported but is also influenced by shared generational exposures (1,19). In this paper we offer and assess an alternative approach to incorporating age in need-based models that is conceptually appealing and has the potential to improve the accuracy and applicability of need-based methods.

In most need-based approaches, age is treated as a typical need indicator. Need indicators are 'legitimate' determinants of healthcare use, and in addition to age, typically include sex and measures of health and morbidity. Age is included and treated in most need-based approaches in the same way as other need indicators, under the implicit assumption that there is a constant relationship between age and observed healthcare use and

morbidity or mortality across different populations (39). However, assuming a uniform relationship between age and healthcare use is neither intuitive nor accurate. There is a wealth of evidence to suggest that different populations have different age-trajectories of morbidity and mortality, which reflect life course exposures and chronic disease patterns (56–60). For example, Canadian Aboriginal and American Native communities suffer a disproportionately greater burden of chronic and acute illnesses and tend to experience these diseases at an earlier age than their non-Aboriginal counterparts (40–42). The assumption that the relationship between age and healthcare use is constant across regions would systematically bias against areas where morbidity and mortality are concentrated at younger ages.

Therefore, we propose a fundamentally different way of treating age in modeling healthcare use. Rather than treating age as a need indicator we have refocused modeling efforts to predicting the age-trajectories of healthcare use. Below we explain in detail how age is currently treated in the need-based approaches and then present an alternative way to model age. We then empirically apply this alternative approach to an administrative dataset of inpatient hospital utilization.

4.1.1 The Current Approach to Incorporating Age

The need-based approach models healthcare utilization by need and non-need factors. Need variables reflect 'legitimate' indicators of healthcare use, and commonly include age, sex, and health status. Non-need variables are factors that influence utilization but are not considered as need for healthcare (10). These often include access to healthcare or wait times. Currently, there is no consensus on which need and non-need variables should be included in need-based approaches (2). Regardless of the number of variables included in these models, age is almost always included as a need indicator for healthcare utilization. Formally, the common modeling of healthcare use is expressed as follows:

$$y_i = \beta_1 + \beta_2 A_i + \beta_3 X_i + \beta_4 Z_i + \varepsilon_i \quad (1)$$

Where y_i is the healthcare use of individual i; β_1 is the coefficient for the intercept; β_2 is a vector of coefficient for a vector of age variables, A_i; β_3 and X_i are, respectively, vectors of coefficients and variables for need indicators; β_4 and Z_i are, respectively, vectors of coefficients and variables for non-need indicators; and ε_i is an error term. Equation (1) is commonly stratified by sex, or alternatively, sex is included in X_i as a need indicator.

Need-expected utilization ($\hat{y}_i^*$) is then obtained by purging equation (1) of the effects of the non-need factors, by holding the values of the non-need factors constant (typically by setting them at their means).

$$\hat{y}_i^* = \hat{\beta}_1 + \hat{\beta}_2 A_i + \hat{\beta}_3 X_i + \hat{\beta}_4 \bar{Z}_i \quad (2)$$

In this way, need factors alone influence the modeling predictions. Need-expected utilization from (2) applies to individuals, and need-expected utilization for a population is obtained by aggregating individual-level values of $\hat{y}_i^*$ according to the population's distribution of need variables.

4.1.2 Consideration of Age: An Alternative Approach

Rather than including age as a need-indicator in the need-based model ($\hat{\beta}_2 A_i$ in equation (2)), we propose an alternative approach that models the age-trajectories of healthcare use. This alternative approach is based on the observation that the age-trajectories of healthcare use typically follow variable but predictable patterns in different populations, and consideration of this predictable variability might allow more appropriate modeling of health care needs. A long and extensive history of epidemiological and demographic research has documented that populations' age-trajectories of morbidity and mortality result from changes in life-course exposures to determinants of health and consequent shifts in chronic disease patterns and cause of death (56,61,64). It is widely recognized that, over the past century, societal advancements have modified the patterns of health and disease. While the average life expectancy has also increased, the maximum life span

21

has remained fairly stable (61,62). Moreover, it appears that healthy behavior changes, such as decreased smoking rates, helped to postpone the onset of chronic diseases until later in life. This phenomenon, known as the compression of morbidity, has resulted in a 'rectangularization' of morbidity curves (56,63). Importantly, this compression of morbidity is not uniform across populations with different rates of health behaviors. Disability increases at much younger ages for some, unhealthier, populations (64). We thus hypothesize that the association between age and the utilization of health services should follow regular, but variable trajectories that reflect the incidence and progression of diseases and disabilities in populations.

Rather than modeling age as one of a series of need indicators, our proposed approach seeks to predict the age-trajectories of healthcare use by region. This can be accomplished by using a growth curve model (65,66). To do this, we begin by incorporating different age-trajectories using a simple growth curve model:

$$y_{ij} = \left(\beta_1 + \delta_{1j}\right) + \left(\beta_2 + \delta_{2j}\right)A_{ij} + \varepsilon_{ij} \text{ (3a)}$$

In this model, the average intercept (β_1) and slope (β_2) include random terms (δ_{1j}) and (δ_{2j}), respectively, to reflect differences between regions; thus acknowledging that age-trajectories are a function of a random intercept $\left(\beta_1 + \delta_{1j}\right)$ and random slope $\left(\beta_2 + \delta_{2j}\right)$ which vary by region. As in equation (2), we can extend equation (3a) to include need and non-need indicators, which predict the regional age-trajectories in the form:

$$y_{ij} = \left(\beta_1 + \delta_{1j}\right) + \left(\beta_2 + \delta_{2j}\right)A_{ij} + \beta_3 X_{ij} + \beta_4 X_{ij}A_{ij} + \beta_5 Z_{ij} + \varepsilon_{ij} \text{ (3b)}$$

Which simplifies to:

$$y_{ij} = \left(\beta_1 + \beta_3 X_{ij} + \delta_{1j}\right) + \left(\beta_2 + \beta_4 X_{ij} + \delta_{2j}\right)A_{ij} + \beta_5 Z_{ij} + \varepsilon_{ij} \text{ (3c)}$$

The regional intercept $(\beta_1 + \beta_3 X_{ij} + \delta_{1j})$ and slope $(\beta_2 + \beta_4 X_{ij} + \delta_{2j})$ describing the age-trajectories are variant between regions, and are modified by need-indicators, as well as non-need indicators $(\beta_5 Z_{ij})$. It would be possible to further allow age-trajectories to vary by non-need factors, but we have not done that here.

Equations (3a), (3b), and (3c) assume that the effect of age on utilization varies between j regions, through the inclusion of a random component for the intercept (δ_{1j}) and slope (δ_{2j}). In equations (3a) to (3c), the age-trajectories are expressed in simple linear form, but this linear transformation of non-linear regression models could also be used according to the same framework.

Unlike a typical need-based approach, our proposed approach models regional age-trajectories of healthcare use, and the other need indicators are used to predict variation in the age-trajectories. In addition, the random coefficients for age capture unexplained variation in the age-trajectories of healthcare use, while providing smoothed estimates of healthcare use by age.

Growth curve modeling is an established method that has been widely employed to examine a range of phenomena such as child growth trajectories (46,67–69), bacterial growth (70–72), and tumor development (47,48,73,74). Useful integration of the growth curve modeling into need-based modeling requires several considerations. We considered the following three questions to be the most important. First, can the different age-trajectories of healthcare use in different regions be modeled with a common functional form? Second, is there sufficient regional variation in the age-trajectories of healthcare to warrant incorporating their variability into modeling? Finally, can ecological-level variables indicating need help predict variation in the regional age-trajectories of healthcare use?

4.2 Methods

4.2.1 Methods Overview

To illustrate and examine the feasibility of modeling the age-trajectories of healthcare use, we modeled the age-trajectories of the rate of inpatient hospital utilization across Canadian census divisions. We employed random coefficient models, also known as hierarchical or mixed models, to estimate growth curve parameters.

The study had three objectives and three corresponding analytical steps. The first objective was to determine whether the age-trajectories of the rate of inpatient hospital utilization could be modeled with a common functional form. To do this, we graphically analyzed the regional age-trajectories of the rate of inpatient hospital use for a select and diverse set of census divisions, and assessed the appropriateness of a variety of functional forms. The second objective was to determine whether there was regional variability in the age-trajectories of the rate of inpatient hospital utilization. The variation in the age-trajectories of the rate of inpatient hospital utilization was described using summary statistics. Then we estimated growth curve models to examine the variation in age-trajectories. The unit of analysis was single year of age by census division, stratified by sex. The dependent variable, the per capita rate of inpatient hospital days of care for census divisions, by single year of age and sex, was computed using Canadian hospital discharge abstract data and census population estimates. The third objective was to determine whether the variability in the age-trajectories of the rate of inpatient hospital utilization was, in part, predictable. We added census division-level variables to the growth curve models, and included cross-level interactions with age to assess whether the age-trajectories of the rate of inpatient hospital utilization could be partially explained by census division-level need indicators.

4.2.2 Data Sources

4.2.2.1 The Discharge Abstract Database (DAD)

We used inpatient hospital utilization data as a measure of healthcare use. We obtained inpatient hospital utilization data from the Discharge Abstract Database (DAD) through the Canadian Institute for Health Information (CIHI). All Canadian hospitals, except those in Quebec, submit their hospital discharge data and day surgery data directly to the CIHI in a standardized way. The DAD includes information regarding transfers; length-of-stay; patient deaths; administrative, clinical, and demographic information; and day surgeries (e.g., cataract patients). Outpatient procedures (e.g., blood samples or x-rays) and all abortion procedures are excluded from the DAD, and were thus not included in the study. Visits to the emergency department that do not result in hospital admittance were also excluded from the DAD. We excluded long-stay hospital visits (those over 60 days) because we were interested in measuring inpatient hospital use for acute care, and most patients who remain in hospital for longer than 60 days are using the hospital as a long-term care facility. We excluded individuals younger than 40 because to assess the feasibility of the approach we wished to examine the rates of hospital use attributed to chronic disease, and the indicator of admissions under the age of 40 are significantly different (influenced by childbearing, accidental injury, and acute illness). We excluded individuals past age 89 because the age-specific population data were unavailable at older ages, such that we were unable to transform healthcare utilization into per capita rates starting at age 90. Thus, the DAD provided the total number of inpatient hospital days, by sex, over three years: 2005, 2006, 2007.

Census divisions located in the territories were not included because the relationship between healthcare utilization and age are assumed to present differently in these areas (75–77). Census divisions in Quebec were not included because the discharge data from Quebec is not retrieved or inputted with the same methodology as the other provinces. With these exclusions, this study included 184 census divisions.

4.2.2.2 The 2006 Census

We obtained population size estimates by age and sex to estimate rates of inpatient hospital use, and ecological-level variables indicating healthcare need for census divisions from the 2006 Census. One fifth of Canadians received the 2006 long-form census by mail and had the option of returning the form via mail, or filling it out and

submitting it electronically (online at Statistics Canada). The 2006 census provides information pertaining to population demographics (population count, age, sex), lifestyle (family make-up, employment), and socioeconomic status (income, education). Census tabulations for 2005, 2006, and 2007 provided summary tabulations of demographic and socioeconomic attributes of census division populations by age and sex, and were accessed through a Canadian Census Analysis provided by the Data Liberation Initiative. Using the DAD inpatient hospital data along with census population information we calculated the average per capita rate of hospital days by census division.

4.2.3 Variables

To represent healthcare utilization we used the rate of per capita inpatient days for each census division by sex, averaged over three years. We included individuals between the ages of 40 and 89, with age measured in single years of age.

We included census division variables in the third step of the analysis in order to predict the variation in the regional age-trajectories of the rate of inpatient hospital use. The 2006 Census provides a range of variables, many of which would not be applicable to our study. To select census variables, we identified the variables that we expected to be the most strongly linked to healthcare utilization. We categorized these possible variables as being related to: income, employment, education, marital status, language, ethnicity, citizenship, immigration, and aboriginal identity. We then selected one or two variables from within each category to be included in the model. In many cases, we excluded a possible variable to avoid conceptual ambiguity or collinearly (e.g., one variable provided the proportion of the population who were Canadian citizens, while another variable provided the proportion of the population who were not Canadian citizens.) Additionally, while many of the identified census variables were unique, some were not conceptually distinct (for example, 'aboriginal identity' and 'aboriginal ancestry' are intuitively associated). In situations where two of the variables were strongly correlated with each other, we chose the one we thought would be most reliable or pertinent to healthcare use. In cases where two variables were conceptually related to each other, but neither one was

intuitively more reliable, we chose the variable with the greatest heterogeneity across census divisions. The variables selected and their definitions are shown in Table 1.

4.2.4 Analysis

Step 1: Determine Whether the Age-Trajectories of Inpatient Hospital Use Could Be Modeled with a Common Functional Form

To identify candidate functional forms, we started by graphically evaluating the relationship between the per capita rate of inpatient healthcare use and age across census divisions. We limited the possible functional forms to: those that were simple with two parameters or less; and those that could be linearly transformed. Functional forms that met these criteria were evaluated for a diverse and select subset of census divisions that included a range of different age patterns. We ran individual regression models for these selected census divisions and assessed the fit. To minimize bias resulting from small age cohorts, we weighted the dependent variable by the relative size of the age-specific population, as compared to the population size for each census division (78). We evaluated the residual variation, and selected the functional form that generated the least systematic error by age. In other words, we chose a functional form based on both the goodness of fit as well as the distribution of the residual variation. Once we had identified a preferred functional form we evaluated the goodness of fit, and variation of the residuals, for all census divisions, stratified by sex.

Step 2: Determine Whether There is Regional Variation in the Age-Trajectories of Inpatient Use

To determine whether the age-trajectories of the rate of inpatient hospital use vary systematically across census divisions, we conducted two types of analysis: (1) descriptive analysis using the Gini coefficient (79–81), and (2) growth-curve modeling of the rate of inpatient use by age, as estimated by random coefficient models.

To complete the descriptive analysis we plotted the cumulative per capita rates of inpatient use (y-axis) according to age, from 40 to 89 (x-axis) for each census division and sex. This is the accumulated use at each age that a person would experience if they experienced the average age-specific rates, and survived from ages 40-89. To summarize age-trajectories of the rate of inpatient hospital use for each census division as a single number, we computed the Gini coefficient. The Gini coefficient is a measure that is widely used to describe inequality. For our purposes, it provides a convenient summary measure of the shape of the age-trajectories curve. Specifically, the Gini Coefficient, which ranges from 0 to 1, was a summary of the concentration of healthcare use by age, and defines the degree to which inpatient use is concentrated at the oldest ages, with 0 representing that healthcare use is consumed equally across all ages and 1 representing that healthcare use is concentrated at the oldest age. Higher Gini coefficients identify census divisions with more curved age-trajectories, where use is concentrated at the older ages, consistent with high compression of morbidity. With the Gini coefficients, we also identified a set of diverse census divisions for assessing the versatility of different functional forms in our objective 1 analysis.

Once we identified an appropriate functional form, we estimated growth curve models, stratified by sex, of the age-trajectories of the rate of inpatient hospital use. We estimated growth curve models by random coefficient models. The dependent variable was the log-transformed rate of inpatient use. The independent variables included age and age-squared. For ease of interpretation of the modeling results, we set zero to age 65 (the rounded median). Additionally, we continued to weight each age-specific rate of healthcare use by the proportion of the total population at each age group.

We ran a series of increasingly complex growth curve models. To work with the data available, we used regionally aggregated data instead of individual level data (shown in equation [3]). The first model (4a) treated the intercept and the slope of the regression as fixed effects:

$$ln\ y_{aj} = \beta_1 + \beta_2 A_{aj} + \beta_3 A_{aj}^2 + \varepsilon_{ij} \quad (4a)$$

As in equation 1, the intercept β_1 was fixed. β_2 is the coefficient for age, A_{aj}, and β_3 is the coefficient for age-squared $A_{aj}{}^2$.

The second model (4b), the random effect model, added a random intercept and slope:

$$\ln y_{aj} = (\beta_1 + \delta_{1j}) + (\beta_2 + \delta_{2j})A_{aj} + (\beta_3 + \delta_{3j})A_{aj}{}^2 + \varepsilon_{ij} \quad (4b)$$

We used likelihood ratio tests to examine whether the introduction of additional random effects in these models significantly improved model fit. To further evaluate the models, we obtained empirical Bayes predictions of age-trajectories of the rate of inpatient hospital use for census division, and conducted residual diagnostics to examine the quality of the predictions across diverse census divisions.

Step 3: Determine Whether the Regional Variation in the Age-Trajectories Could Be Predicted

In the final step of the analysis, we examined whether the census division-level variables could predict regional variation in the age-trajectories of the rate of inpatient hospital use. We added census division variables as fixed effects into the growth curve model from Step 2:

$$\ln y_{aj} = (\beta_1 + \beta_4 X_j + \delta_{1j}) + (\beta_2 + \beta_5 X_j + \delta_{2j})A_{aj} + (\beta_3 + \delta_{3j})A_{aj}{}^2 + \varepsilon_{ij} \,(4c)$$

In addition, not shown in the equations, we allowed our coefficients to be correlated. One could further complicate model (4c) by allowing non-need variables to modify the coefficients, however, it was not explored in this paper.

Table 1 describes the census variables we included in our analysis. For ease of interpretation, we centered the census division variables at their individual means. This model retained the intercept and the slope of age and age-squared as random effects. We also included cross-level interaction terms between each census variable and age to allow

the census variables to predict the slope, in addition to the intercept. While interactions with age-squared were also explored, they were not included in presented models as they offered minimal improvements in model fit at the expense of considerable complexity. To examine whether the introduction of census division variables improved the model fit we performed a likelihood ratio test. In addition, we obtained empirical Bayes predictions and used them to graphically assess the fit of model predictions for diverse census divisions.

4.3. Results

Figure 1 shows the empirical variation in age-specific rates of inpatient hospital use. There was little variation across census divisions in the age-specific rate of healthcare use prior to age 50, and generally, the most variation across census divisions was seen in the healthcare use after age 75, and continued to increase with age.

With respect to the first objective, our results showed that a simple two-parameter polynomial, predicting the log of the rate of inpatient use as a function of age and age-squared, could be used to model the age-trajectories of the rate of inpatient hospital use for census divisions displaying a diversity of age trajectories. The second order polynomial did not generate much systematic error when fit across the Canadian census divisions included in the study. That is, the residual variation was not concentrated at any specific age range, and was evenly distributed around zero. The lack of systematic error was confirmed when evaluating the fit of the model individually to a set of diverse census divisions. Largely, the residual variation was present at older ages, but was generally not systematically concentrated at any age. The second order polynomial provided the best fit for census divisions with the largest population sizes, presumably because the data were not as susceptible to random error. Census divisions with small population sizes were the least likely to be successfully modeled with the two-parameter polynomial equation, as rates fluctuated with age, especially at the older ages where population counts were small.

In addition to the second order polynomial, we tested the fit of other functional forms. The Gompertz functional form fit the data with reasonable precision, but was too complex for the purposes of this study. A simple linear equation had a great deal of systematic error at the older ages, across all census divisions. The logistic functional form did not fit the data at the oldest ages.

Analysis based on both the Gini coefficients and the growth curves showed that there was considerable variability in the age-trajectories of the rate of inpatient use for both sexes. Figure 1 describes the variability in census divisions-level age-specific rate. Gini coefficients for census divisions ranged from 0.38 to 0.76 for females and from 0.41 to 0.71 for males. Figure 2 shows the average rate of inpatient use by age for census divisions that were below the 10th percentile, between the 10th to 25th, 25th to 50th, 50th to 75th, 75th to 90th percentiles, and above the 90th percentiles of the Gini coefficient. For females and males, the shape and the distribution of the age-trajectories of the rate of inpatient hospital use varied most significantly after age 70. The exception to this was seen in males in census divisions that ranged from the 10th to 25th percentile of their Gini coefficients. The relationship between healthcare use and age for this cohort presented differently from the other cohorts after age 48. For this group, inpatient use is less associated with age. Additionally, prior to age 61 for females and 66 for males there was very little observable variation in the healthcare use by age across census divisions, and variation only occurs after age 71 for females and 70 for males.

We confirmed these results through the growth curve analysis. Table 2 shows results from the growth curve analyses. Models 1 and 2 in Table 2 show estimates of the fixed and random coefficient models based on a second order polynomial for age regressed on the log of the rate of inpatient use. For both sexes the inclusion of random coefficients for the intercept, age and age-squared significantly improved model fit relative to a fixed effect model ($p<0.001$). While not shown in the table, we also found that the inclusion of a random intercept alone ($p<0.001$), then the inclusion of a random coefficient for age ($p<0.001$) and age-squared ($p=0.03$) each resulted in incremental improvements in model fit. A model with random intercept alone revealed that 50% of the variation in the rate of

inpatient use for females and 38% of the variation in the rate of inpatient use for males was explained by between-census division variation. The standard deviations of all three random coefficients in Model 2 are substantial, relative to the magnitude of the fixed effect coefficients, which is an indication of substantial variation in the age-trajectories of the rate of inpatient hospital use. The negative correlation between the random coefficients for the intercept and age is also noteworthy, and consistent with observations from Figure 2. Census divisions that have flatter age-trajectories of the rate of inpatient hospital use at age 65 (smaller slope) tend to have higher rate of inpatient use at age 65 (higher intercept). Taken together, the Gini analysis and the growth curve models confirmed the second objective of the study and showed that there was significant variability in the regional age-trajectories of the rate of inpatient hospital use.

Census division-level variables were predictive of both the level and age-trajectories of the rate of inpatient use. Our analysis found that the introduction of census division-level variables into the random coefficient model significantly improved the fit of the model for both sexes, as seen through a likelihood ratio test (p<0.001). Main effects of family variables (proportion divorced and proportion lone-parent families), socioeconomic variables (education, employment rate, and employment participation rate), and ethnic variables (proportion of Canadian citizens, and proportion of first generation immigrants) were among the significant variables, and helped to explain the census division variation in the intercept for both sexes (Table 2, Model 3). The introduction of these variables reduced the standard deviation of the intercept (corresponding to age 65) from .28 to .15 for females, and .24 to .12 for males. Interactions between census division-level variables and age were also explored to examine whether the slope could be partially explained, and those that were significant were retained in Model 3. Table 2 also shows that for both sexes, the interactions between age and proportion of the population that was divorced, and age and the median family income were found to be significantly associated with the slope. For females, the proportion of lone parent families, and for males the proportion of people who identify as Aboriginal, were also significant. To reduce complexity, and because they made only small contributions to model fit, interaction terms between age-squared and the census division-level variables were not included in the final model.

4.4. Discussion

An association between age and healthcare use is well established. This study focused on this relationship and proposed to model the age-trajectories of healthcare utilization rather than including age as an indicator of need for healthcare utilization. Using standardized national level administrative inpatient hospital use and census data, we showed the feasibility and promise of this new approach. This study had three significant findings. First, the age-trajectories of the rate of inpatient hospital use, after age 40, exhibited variable but predictable patterns across regions, and these patterns could be reasonably modeled using a second order polynomial. Second, there was significant regional variation across the age-trajectories of the rate of inpatient hospital use. Third, census division-level variables could help predict this regional variation in age trajectories of the rate of inpatient hospital use. Taken together, our study indicated that modeling the age-trajectories of healthcare utilization, rather than simply including age as a need-indicator, can improve need-based approaches.

The methods proposed in this paper are conceptually aligned with current knowledge of the relationship between age and morbidity. Additionally, there are a number of reasons why the modeling the age-trajectories of healthcare utilization has the potential to improve the accuracy, fairness, and versatility of need-based modeling approaches and healthcare.

Modeling age-trajectories of healthcare utilization is theoretically intuitive, and consistent with what has been learned through decades of epidemiologic and demographic research on mortality and morbidity. It has been widely recognized that, over the past century, societal advancements and changes in life-course exposure have altered the age-trajectories of health and disease (61,63). In most western countries, including Canada, there has been a transition from a disease pattern produced by infectious diseases, with high mortality concentrated at younger ages, to a disease pattern produced by chronic diseases, concentrated at older ages. Orman, who was among the first to recognize the shift in the patterns of disease (the 'epidemiological transition theory'), attributed the

early decline in mortality and morbidity to a complex array of factors associated with 'modernization' (62), including: developments in health technologies, advancements in public health, and improvements in the social determinants of health.

A separate, but consistent observation, the compression of morbidity, suggests that this modernization will also shift the age-trajectories of disease to the right, becoming more 'rectangluarized' (13,24). Thus, while the average life expectancy has increased significantly in western countries, the maximum life span has not increased proportionally, and in fact appears to remain relatively stable. This suggests that improvements in health have not significantly modified the intrinsic human life span (82), rather, have reduced premature mortality and morbidity (61,62). The use of regional age-trajectories in healthcare use modeling may better capture these high order epidemiological observations and indeed our findings support this prediction. The regional age-trajectories of the rate of inpatient hospital use followed a predictable pattern, and census division-level variables helped explain the regional variability, consistent with the compression of morbidity and epidemiologic transition theories. Additionally, modeling the age-trajectories of morbidity is an intuitive extension of the widely used technique of modeling the age-trajectories of mortality, which has been widely used for decades to develop 'model life tables' for purposes of estimating mortality patterns in populations with deficient vital statistics data (83–85).

The proposed method may improve the accuracy of estimates of need-based resource requirements by exploiting information on the patterning of age-trajectories of healthcare use. By employing information on the functional form of the age-trajectories of healthcare use, more smooth and precise estimates can be made for small areas. Moreover, the methods proposed are amenable to Bayesian methods to predict the age-trajectories of healthcare use for small areas, borrowing power from other areas with similar attributes. Such methods are widely used for purposes of small area estimation (49,86–89). Accordingly, the methods we propose offer the potential to extend need-based methods to smaller areas of geography than are feasible with current methods.

The proposed method has the potential to provide for a more equitable and fair distribution of healthcare resources, as it better captures heterogeneity in the need for healthcare services by age. Current models, which incorporate age as a fixed effect, assume that average healthcare costs by age are equally applicable to different populations. As such, they underestimate healthcare requirements for populations in which needs are shifted to younger ages as a result of higher incidence and earlier onset of chronic disease. Examples of such populations are Canadian Aboriginal and American Natives, who suffer disproportionately from a number of chronic and acute illnesses, and tend to experience these diseases at an earlier age than their non-Aboriginal, non-Native counterparts (40–42).

The proposed approach can be extended to both a variety of data and to commonly used need-based modeling approaches. The need-based model in our analysis is fairly rudimentary. It employed only ecological data, examined only one type of healthcare utilization (the rate of inpatient hospital utilization), and only examined ages over age 40. As well, the census division variables used as estimates of need were limited, and no effort was made to adjust the models for non-need indicators of utilization. However, the method we propose here could readily be extended to individual level data, clustered by geography, and both individual and ecological need and non-need indicators of utilization could be included in the models. As is customary, need–expected utilization could still be obtained by purging predicted utilization of the influence of non-need indicators. Extension of the approach we propose to other types of healthcare utilization and age ranges will likely require different functional forms. Splines offer a promising approach where different functional forms are appropriate for different ranges of age.

To meet the third objective of this study we needed to show that the variability in the regional age-trajectories of the rate of inpatient hospital use could be partially explained. We therefore selected a limited set of variables to demonstrate viability. Future work should carefully identify variables based on the conceptual association between the variable and healthcare use within the constraint of the data availability.

Our proposed approach could have research and policy implications that extend across a number of sectors in health services. For example, these results may be pertinent in resource allocation methodology. Using the age-trajectories of healthcare use is an attractive alternative to current need-based resource allocation models because it is conceptually simple. Alternatively, this approach could benefit human resource planning, emphasizing the effects of demographic changes on the health human resource needs of a population, and applying the population's demographic mix to population-utilization ratios, determining planning requirements (90).

This study has a number of limitations. For practical application of the proposed method, further consideration of the appropriate level of geographical analysis is necessary. First, we found that the age-trajectories of the rate of inpatient hospital utilization were sufficiently variable at the level of the census division. However, an appropriate level of geography to be employed in the proposed method may well be different depending on types of healthcare. The existing funding structure may also be an important consideration in determining an appropriate level of geography. In Canada, for example, need-based resource allocation models are typically used to allocate resources from the province to health regions, and viability of the proposed approach at the provincial level is an open question.

A second limitation was that the data we received from CIHI contained hospital in-patients days by year up to age 95. However, we were only able to obtain population counts from Statistics Canada up to age 90. The first objective of our project was to find a functional form with which to model the data. To minimize the residual error that would have resulted from aggregating ages 90 and above, we chose to exclude these patients and limit our analysis to people 89 and under. This is unlikely to have had a significant impact on the results, as the pattern of inpatient hospital use is well established by age 89. Additionally, this study concluded that a second order polynomial could model the age-trajectories of healthcare usebetween the ages of 40 and 89. However, it would be unlikely that the same functional form would fit the data for ages under 40. Additionally,

future work would have to confirm that the polynomial functional form would fit to other types of healthcare services use.

Third, we were unable to include those census divisions located in Quebec or the territories in this study. Additionally, this data set did not contain long-term stay, day surgeries, or emergency visits that did not result in a hospital admission. To expand this methodology these exclusions would have to be explored in future research.

The fourth limitation was that, by using the rate of inpatient hospital use to reflect healthcare use, this study did not acknowledge patient acuity. It is possible that two regions may have the same per capita rate of inpatient healthcare use, but different levels of resource use by virtue of the case-severity. Also, this project does not address patient death. That is, some regions may have lower hospital inpatient hospital use because the patients from that region have a higher mortality rate (lower chance of survival) than in another region.

4.5. Conclusion

Need-based approaches can be refined and improved by changing the way age is incorporated into utilization models. In this paper, we found that age was an integral factor in predicting healthcare utilization; and modeling the age-trajectories of health care utilization, rather than including age as a standard need indicator, enhanced need-based approaches. Incorporating the age-trajectories of healthcare utilization may reduce regional variability in healthcare modeling and may permit a more appropriate allocation of resources, especially to communities with early onset of morbidity.

We found that this methodology is feasible and has the potential to work in a Canadian setting. This methodology is consistent with the life-course perspective: there are regular patterns of mortality and it is intuitive to extend this logic to morbidity. The amendment to need-based approaches we proposed in this paper – modeling the age-trajectories of

healthcare use – could ameliorate the approach and address some of the current limitations, thus widening its potential for application.

Table 1: Description of Census Division-Level Variables

Classification	Variable	Description	Mean (Standard deviation)
Income	Family income *(Fam income)*	Average family after-tax income (in thousands)	59.73 (10.41)
Employment	Employment participation rate *(Em part rate)*	The size of the labor force expressed as a percentage of the total population 25 years of age and over	65.1 (6.51)
	Employment rate *(Em rate)*	The number of persons employed expressed as a percentage of the total population 25 years of age and over	60.8 (8.22)
Education	High school educational attainment *(HS educ)*	Proportion of the population 25 to 64 years of age with a high school diploma or equivalent	25.9 (3.73)
Marital Status	Divorced *(divorced)*	Proportion of the population who have been divorced and are currently not married	5.75 (1.36)
Lone parent households	Proportion of the population who are living as lone-parents *(ln parent)*	A mother or a father, with no spouse or common-law partner present, living in a dwelling with one or more children.	14.5 (3.73)
Language and Ethnicity	Visible minority *(vis minor)*	Number of persons who self-identify as a visible minority	4.12 (7.24)
Citizenship	Canadian citizenship *(Cdn citizens)*	The number of persons who, at the time of the Census, have official Canadian citizenship	97.6 (2.18)
Dwelling characteristics	Regular repairs required *(%home repair)*	The percentage of respondents, who, in the opinion of the respondent, state that their dwelling requires regular repairs	58.9 (8.14)
Generational status	First generation Canadian *(1st gen)*	Percentage of people who were born in Canada, but neither of whose parents were born there	10.4 (9.51)
Aboriginal Status	Aboriginal Identity *(Aborig iden)*	Proportion of the population with an affiliation with an Aboriginal group that is North American Indian, Métis or Inuit	8.72 (11.90)

Figure 1: Relationship between healthcare use and age
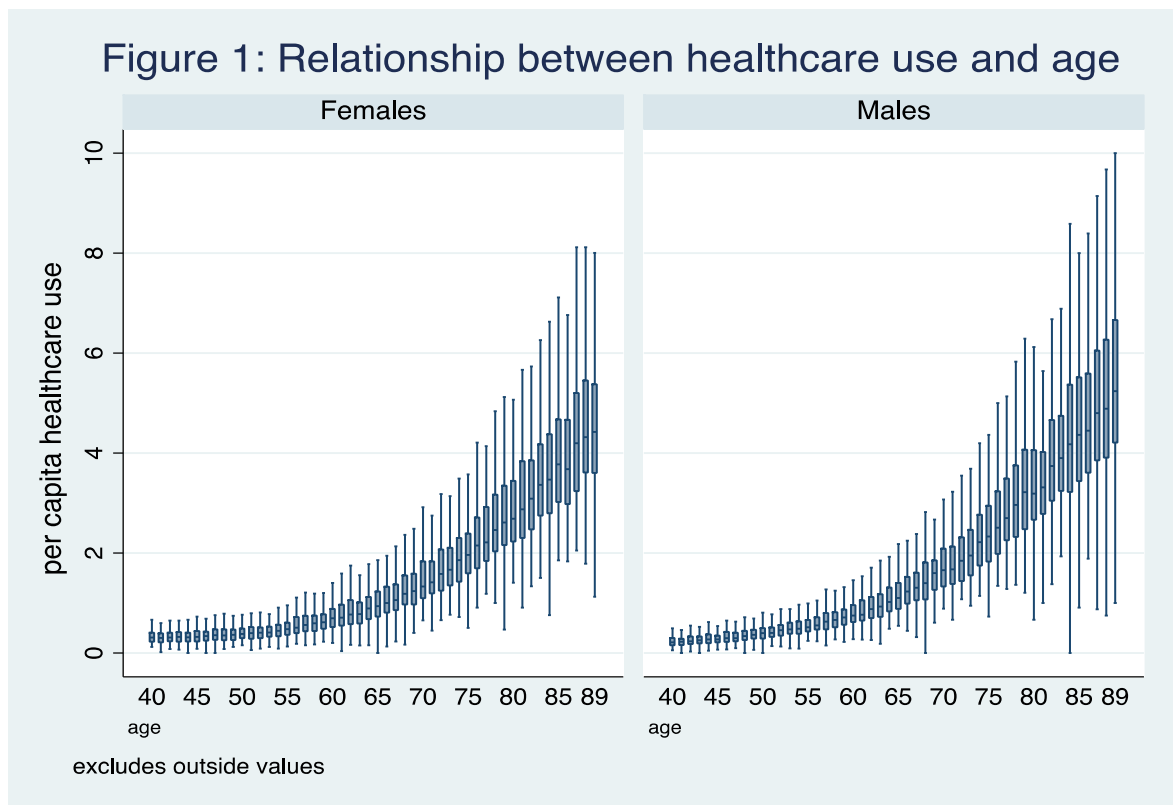
Figure 1: A Box Plot of the Relationship between Observed Inpatient Hospital Use, By Sex

Note: The thick grey bars in the center of the vertical lines represent the 25th to 75th percentiles of the per capita rate of observed inpatient use. The thin grey vertical lines stemming above represent the 75th to 90th percentiles, and the thin grey vertical lines stemming below represent the 25th to 10th percentiles. The extreme values, below the 10th percentile and above the 90th percentile, are not shown. In total these graphs show the three-year aggregated data, as provided by the DAD, for 184 census divisions.
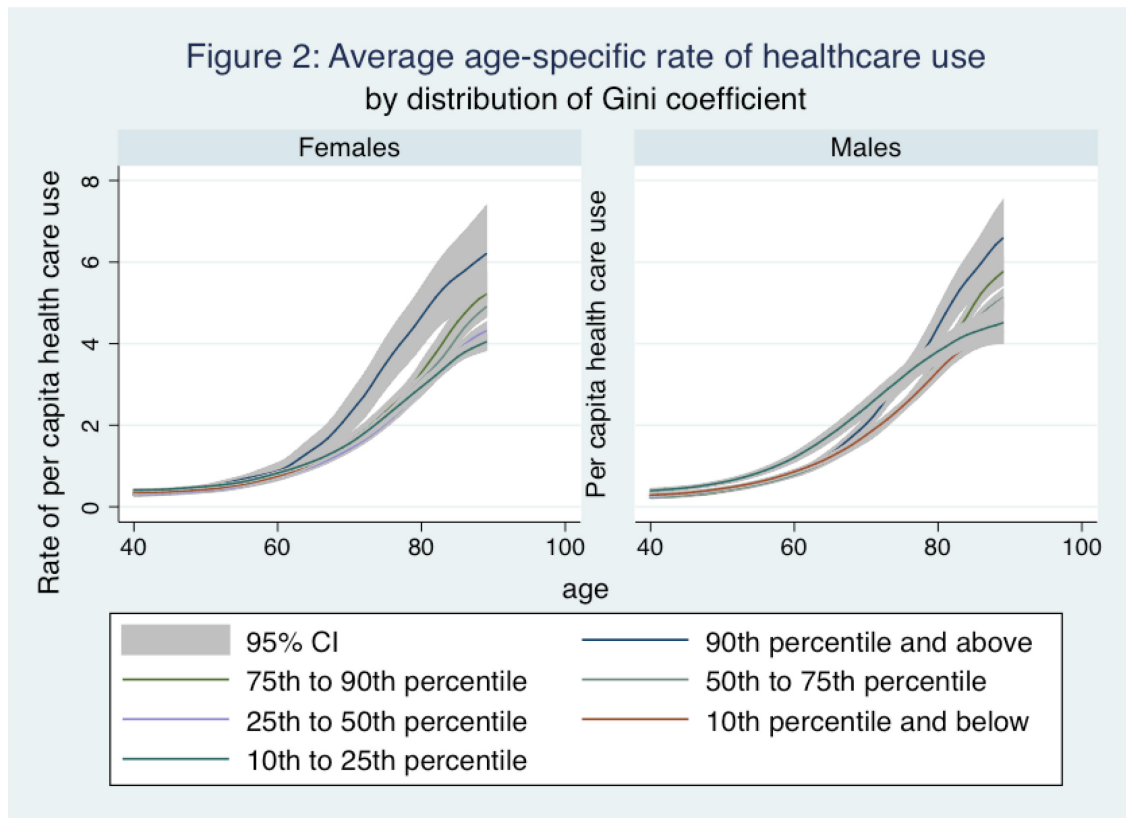
Figure 2: The Average Age-Specific Rate of Healthcare Use by Distribution of Gini Coefficient

Note: The Gini coefficient of one in our data means that all healthcare use is concentrated at the oldest age group where a Gini coefficient of zero indicates that healthcare use is spread evenly across all ages. We estimated the Gini coefficient for each census division and sex and, based on the distribution of the Gini coefficients across census divisions by sex, classified them into six groups for men and women separately: (1) the bottom 10th percentile, (2) 10th-25th percentiles, (3) 25th-50th percentiles, (4) 50th-75th percentiles, (5) 75th-90th percentiles, and (6) the top 10th percentile. We then graphed the three-year average relationship between the rate of inpatient healthcare use (received from the DAD) and age for each of these six groups, for men and women separately.

Table 2: The Results of the Growth Curve Analysis as Estimated by Random Coefficient Models

Parameter	Model 1		Model 2		Model 3	
Fixed Component	**Females**	**Males**	**Females**	**Males**	**Females**	**Males**
Intercept	-.07(.02)	.11 (.017)	-.07 (.02)	.11 (.19)	-.05 (.012)	.13 (.01)
Age	.06 (.0002)	.07 (.0002)	.06 (.0004)	.07 (.0003)	.06 (.0003)	.07 (.0004)
Age-squared	.0004 (.00002)	-.0004 (.00002)	.0004 (.00002)	-.0004 (.00003)	.0004 (.00002)	-.0004 (.00003)
Divorced					-.03(.012)*	-.032(.01)*
Ln parent					.014(.005)*	.013(.004)*
%home repair					-.01(.003)*	-.007 (.003)*
Cdn citizens					-.06(.02)*	-.06(.015)*
Aborig iden					.003(.002)	.001(.002)
1st gen					-.03(.006)*	-.03(.004)*
Em part rate					.03(.007)*	.02(.005)*
Em rate					-.02(.006)*	-.02(.005)*
HS educ					.02(.004)*	.014(.004)*
vis minor					.008(.005)	.007(.004)
Fam income					.006 (.003)*	.003 (.002)
divorced* age					-.001 (.0004)*	-.002 (.0004)*
ln parent*age					-.0005 (.0001)*	
Aborig iden*age						-.0002 (.00006)*
Fam income*age					.001 (.0007)*	.0002 (.00008)*
Random Component						
Sd (intercept)			.28 (.016)	.24 (.014)	.15 (.009)	.12 (.008)
Sd (age)			.005 (.0003)	.005 (.0004)	.004 (.0003)	.004 (0003)
Sd (age-squared)			.0003 (.00002)	.0003 (.00002)	.0003 (.0002)	.0003 (.00002)
Corr(inter, age)			-.24(.08)	-.23(.08)	-.22(.09)	-.30 (.09)
Corr(age, age-squared)			-.20(.10)	-.024 (.10)	-.38(.12)	.011(.011)
Sd (residuals)	.26 (.002)	.28 (.002)	.25 (.002)	.27 (.002)	.24(.002)	.26(.002)

Note: The results of the growth curve analysis for three models. Model 1 is the fixed regression; Model 2 is the random effects regression on the log of the rate of inpatient hospitalization that includes the intercept, age and age-squared as random effects; and model 3 is the full regression on the log of the rate of inpatient hospitalization, that treats that intercept, age and age-squared as random effects and includes census-division level variables as fixed effects.

CHAPTER 5: CONCLUSION

Our study explored the feasibility and usefulness of a need-based approach that models the age-trajectories of the rate of inpatient hospital use. To do so, we proceeded through three analytic phases. First, we examined whether a predictable pattern exists for the relationship between age and inpatient use across census divisions, and if so, whether this relationship could be modeled with a single functional form. We found that the pattern of the age-use relationship was consistently shaped across most census divisions, and that a two-parameter polynomial fits the data with little residual error at any age group. Census divisions with the smallest populations (those in the bottom tenth percentile of population size) were the least likely to follow this pattern.

Second, we examined whether the age-trajectories of the rate of inpatient hospital use varied systematically across census divisions, by performing a descriptive analysis and applying growth curve models. These results confirmed that there was significant variability in the regional age-trajectories of the rate of inpatient hospital use. The majority of the variation was concentrated at ages older than 45 and ages younger than 85. For most census divisions, the average inpatient healthcare utilization remains stable from ages 40 to 45, and thus there is little regional variation at these ages. Most census divisions start to see significant increases in the rate of healthcare use between ages 50 and 60, and it is at these ages that the rates of healthcare use start to vary between census divisions.

Third, we evaluated whether including census division-level variables could help predict the regional variation in the age-trajectories of the rate of inpatient hospital use. We found that census division-level variables were partially predictive of both the level and age-trajectories of the rate of inpatient use. The introduction of census division-level variables into the random coefficient model significantly improved the fit of the models for both sexes.

This study had three significant findings. The first was that the age-trajectories of healthcare use exhibited a predictable pattern across regions, and this pattern could be modeled with reasonable accuracy using a two-parameter polynomial. The second finding was that there was significant regional variation across the age-trajectories of inpatient healthcare utilization. The third finding was that ecological-level data could help to explain some of the regional variability in inpatient healthcare use. Taken together, our study indicates that age should not be treated as a typical need-indicator. Rather, age-trajectories of healthcare utilization vary significantly between regions, and jurisdictions that use need-based resource allocation formulas might benefit from modeling the age-trajectories of healthcare utilization.

The methods proposed in this thesis are conceptually aligned with current knowledge of the relationship between age and morbidity. Additionally, there are a number of reasons why modeling the age-trajectories of the rate of inpatient hospital use has the potential to improve the accuracy, fairness, and versatility of need-based modeling approaches and healthcare.

Modeling age-trajectories of inpatient healthcare utilization is theoretically intuitive, and consistent with what has been learned through decades of epidemiologic and demographic research on mortality and morbidity. The examination of the age-trajectories of the rate of inpatient hospital use in this study is aligned with two established theories of aging and disease: the epidemiological transition and the compression of morbidity. It has been widely recognized that, over the past century, societal advancements and changes in life-course exposure have altered the age-trajectories of health and disease. In most western countries, including Canada, there has been a transition from a disease pattern produced by infectious diseases, with high mortality concentrated at younger ages, to a disease pattern produced by chronic diseases, concentrated at older ages. Orman, who was among the first to recognize the shift in the patterns of disease, ('the epidemiological transition theory'), attributed the early decline in mortality and morbidity to a complex array of factors associated with 'modernization',

including: developments in health technologies, advancements in public health, and improvements in the social determinants of health.

A separate, but consistent observation, 'the compression of morbidity', postulates that this modernization will shift the age-patterns of disease to the right, becoming more 'rectangluarized' (56,64,91). Thus, while the average life expectancy has increased in western countries, the maximum life expectancy appears to remain fixed. This suggests, improvements in health have not modified the human life span itself; rather, they have reduced premature mortality and morbidity (61,62). The use of regional age-trajectories in healthcare modeling may better capture these high order epidemiological observations, and indeed our findings supported this. The regional age-trajectories of the rate of inpatient hospital use followed a predictable pattern, and census division-level variables helped explain regional variability, consistent with the compression of morbidity and epidemiologic transition theories. Additionally, modeling the age-trajectories of morbidity is an intuitive extension of the widely used technique of modeling the age-trajectories of mortality. The age-trajectories of mortality are so extensively relied upon, they form the conceptual basis of life tables.

The proposed method may improve the accuracy of estimates for need-based resource requirements by exploiting information on the patterning of age-trajectories of healthcare use. By employing information on the functional form of the age-trajectories of healthcare use, smooth and precise estimates can be made for small areas. Moreover, the methods proposed are amenable to Bayesian methods to predict the age-trajectories of inpatient healthcare use for small areas, borrowing power from other areas with similar attributes. Such methods are widely used for purposes of small area estimation. Accordingly, the methods we proposed offer the potential to extend need-based methods to smaller areas of geography than are feasible with current methods.

The proposed method has the potential to provide a more equitable and fair allocation of healthcare resources than the current method does, as it better captures heterogeneity in healthcare services use by age. Current models, which incorporate age as a fixed effect,

assume that average healthcare use by age is applicable to different populations. As such, they underestimate healthcare requirements for populations in which needs are shifted to younger ages as a result of higher incidence and earlier onset of chronic disease than the average population, and will unfairly benefit populations where disease incidents are delayed to older ages. For example, Canadian Aboriginal and American Native communities suffer disproportionately more from a number of chronic and acute illnesses, and tend to experience these diseases at an earlier age than their non-Aboriginal counterparts (40–42). In Canada cardiovascular disease is the leading cause of death within both Aboriginal and non-aboriginal Canadians (43). However, age-standardized mortality rates among Aboriginal women are 61% higher compared to non-Aboriginal women. Additionally, the stroke mortality rate is 44% and 93% higher among Aboriginal men and women (respectively) than the general Canadian population (44,45). Modeling the age-trajectories of inpatient healthcare use would rectify some of this inequity by appreciating the differences in the age patterns of disease.

The approach we propose can be extended to a variety of different data and need-based modeling approaches commonly used. The need-based model in our analysis is rudimentary. It employed only ecological data, examined only one type of healthcare utilization (the rate of inpatient hospital utilization), and only examined ages over age 40. As well, the census division variables used as estimates of need were limited, and we did not adjust for non-need indicators of utilization. However, the method we propose could readily be extended to individual level data, clustered by geography, and both individual and ecological indicators of need and non-need indicators of utilization could be included in the models. As is customary, need–expected utilization could still be obtained by purging predicted utilization of the influence of non-need indicators. Extension of the approach we propose to other types of healthcare utilization and age ranges will likely require different functional forms. Splines, a method by which you run regression in connected pieces (78), offer a promising approach where different functional forms are appropriate for different ranges of age.

To meet the third objective of this study, to determine whether ecological-level variables could help predict the regional age-trajectories of the rate of inpatient hospital use, we only needed to prove that the variability in the regional age-trajectories of the rate of inpatient hospital use could be partially explained. We therefore selected a limited set of variables to show viability. Future work should carefully identify the most appropriate variables based on the conceptual association between the variable, healthcare use and data availability.

With the methodological novelty and feasibility as well as data flexibility, our proposed approach could have research and policy implications that extend across a number of sectors in health services. For example, these results may be pertinent in resource allocation methodology. Using the age-trajectories of healthcare use is an attractive alternative to current need-based resource allocation models because it is conceptually simple. The relationship between age and healthcare use is so well understood that this model provides an intuitive approach to understanding resource allocation. The relationship between mortality and age is widely used in a variety of contexts, and is so accepted that it forms the basis of life tables. Extending this logic to the age-trajectories of morbidity could also provide wide spread benefit. Moreover, modeling the age-trajectories of healthcare is more transparent than the current model because it follows a logical pathway. We know that age is a strong proxy for mortality and morbidity, and thus makes a powerful indicator of healthcare use. If we work to explain the variation in the age-trajectories of healthcare use, rather than trying to explain the variation of healthcare need in addition to age, we make the model more conceptually parsimonious.

This study has a number of limitations. For practical application of the proposed method, further consideration of the appropriate level of geographical analysis is necessary. We found that the age-trajectories of the rate of inpatient hospital utilization was a sufficient variable at the level of the census division. However, an appropriate level of geography to be employed in the proposed method may well be different depending on types of healthcare, needing to be adjusted with consideration for the existing funding structure. In Canada, for example, need-based resource allocation models are typically used to allocate

resources from the province to health regions, and viability of the proposed approach at the provincial level is an open question.

The data we received from CIHI contained hospital in-patients days by year up to age 95. However, we were only able to obtain population counts from Statistics Canada up to age 90. The first objective of our project was to find a functional form with which to model the data. To minimize the residual error that would have resulted from aggregating ages 90 and above, we chose to exclude these patients and limit our analysis to people 89 and under. This is unlikely to have had a significant impact on the results, as the pattern of healthcare use is well established by age 89. Additionally, this study concluded that a second order polynomial could model the age-trajectories of inpatient hospital use between the ages of 40 and 89. However, it would be unlikely that the same functional form would fit the data for ages under 40. Additionally, future work would have to confirm that the polynomial functional form would fit to other types of healthcare services use.

Because of the data source we accessed, we were unable to include those census divisions located in Quebec or the territories in this study. Additionally, this data set did not contain long-term stay, day surgeries, or emergency visits that did not result in a hospital admission. To expand this methodology these exclusions would have to be explored in future research.

By using the rate of inpatient hospital use to reflect healthcare use, this study did not acknowledge patient acuity. It is possible that two regions may have the same per capita rate of inpatient healthcare use, but different levels of resource use by virtue of the case-severity. Also, this project does not address patient death. That is, a region may have lower hospital inpatient hospital use because the patients from that region have a higher mortality rate (lower chance of survival) than in another region.

To conclude, our results indicate that age is an integral factor in predicting healthcare utilization, and should not be treated simply as a need-indicator. Current need-based

approaches that include age as a need-indicator assume that the effect of age is invariant between regions, and systematically bias against populations who experience earlier onsets of disease. While future studies are necessary to further validate this method, it appears that modeling the age-trajectories of healthcare need could greatly benefit need-based resource allocation models.

REFERENCES

1. Smith A, (Canada) DU. Allocating for the Dying: Age and End-of-life in Need-based Resource Allocation Models. Dalhousie University (Canada); 2008.

2. Kephart G, Asada Y. Need-based resource allocation: different need indicators, different results? BMC Health Services Research. 2009;9(1):122.

3. Heligman L, Pollard JH. The age pattern of mortality. Journal of the Institute of Actuaries. 1980;107(1):49–80.

4. Caughley G. Mortality patterns in mammals. Ecology. 1966;47(6):906–18.

5. Thatcher AR. The long-term pattern of adult mortality and the highest attained age. Journal of the Royal Statistical Society: Series A (Statistics in Society). 1999;162(1):5–43.

6. Carr-Hill RA, Sheldon TA, Smith P, Martin S, Peacock S, Hardman G. Allocating resources to health authorities: development of method for small area analysis of use of inpatient services. BMJ. 1994 Oct 22;309(6961):1046–9.

7. M B, G M. Equity in Health Care from a Communitarian Standpoint. 2002 Jun [cited 2012 Apr 16]; Available from: http://www.ingentaconnect.com/content/klu/hcan/2002/00000010/00000002/0509 0223

8. Gwatkin DR, Bhuiya A, Victora CG. Making health systems more equitable. The Lancet. 2004;364(9441):1273–80.

9. Mooney G. Inequity in Australian health care: how do we progress from here? Aust N Z J Public Health. 2003;27(3):267–70.

10. Asada Y, Kephart G. Understanding different methodological approaches to measuring inequity in health care. Int J Health Serv. 2011;41(2):195–207.

11. Eagar K, Grant P, Lin V. Health Planning: Australian Perspectives. Allen & Unwin Academic; 2002.

12. Pearson M. Allocating public resources for health : developing pro-poor approaches. DFID. Health systems resource centre (HSRC).

13. Rice N, Smith PC. Ethics and Geographical Equity in Health Care. J Med Ethics. 2001 Aug 1;27(4):256–61.

14. Maynard A, Ludbrook A. Applying resource allocation formulae to constituent parts of the U.K. Lancet. 1980 Jan 12;1(8159):85–7.

15. McIntosh T. Population Health and Health Reform: Needs-Based Funding in Five Provinces. CPSR. 2010 Apr 26;4(1):42–61.

16. Mays N. Geographical resource allocation in the English National Health Service, 1971-1994: the tension between normative and empirical approaches. Int J Epidemiol. 1995;24 Suppl 1:S96–102.

17. Birch S, Chambers S. To each according to need: a community-based approach to allocating health care resources. CMAJ: Canadian Medical Association Journal. 1993 Sep 1;149(5):607.

18. Lake M. Modeling "Need" in health sector resource allocation formulae: Application in a low-income country. Health System Financing.

19. Manton KG, Stallard E, Corder L. Changes in the age dependence of mortality and disability: cohort and other determinants. Demography. 1997 Feb;34(1):135–57.

20. Patrick DL, Erickson P. Health status and health policy. Oxford Univ. Press New York; 1993.

21. Case A, Paxson CH. Sex Differences in Morbidity and Mortality. Demography. 2005;42(2):189–214.

22. Morris S, Sutton M, Gravelle H. Inequity and inequality in the use of health care in England: an empirical investigation. Soc Sci Med. 2005 Mar;60(6):1251–66.

23. Lu JR, Leung GM, Kwon S, Tin KY., Van Doorslaer E, O'Donnell O. Horizontal equity in health care utilization evidence from three high-income Asian economies. Social Science & Medicine. 2007;64(1):199–212.

24. Asthana S, Gibson A, Moon G, Dicker J, Brigham P. The pursuit of equity in NHS resource allocation: should morbidity replace utilisation as the basis for setting health care capitations? Social science & medicine. 2004;58(3):539–51.

25. Steenland K, Henley J, Calle E, Thun M. Individual- and Area-Level Socioeconomic Status Variables as Predictors of Mortality in a Cohort of 179,383 Persons. Am. J. Epidemiol. 2004 Jun 1;159(11):1047–56.

26. Newbold KB, Eyles J, Birch S. Equity in health care: methodological contributions to the analysis of hospital utilization within Canada. Social science & medicine. 1995;40(9):1181–92.

27. Patterson WM. Mortality, morbidity, and resource allocation. British Medical Journal (Clinical research ed.). 1981;282(6258):151–151.

28. Gibson A, Asthana S, Brigham P, Moon G, Dicker J. Geographies of need and the new NHS: methodological issues in the definition and measurement of the health needs of local populations. Health & Place. 2002;8(1):47–60.

29. Asada Y, Kephart G, Hurley J, Yoshida Y, Smith A, Bornstein S. The role of proximity to death in need-based approaches to health care. Health Policy.

30. Wilkinson RG, Marmot MG. Social determinants of health: the solid facts. World Health Organization; 2003.

31. Raphael D. Social determinants of health: Canadian perspectives. Canadian Scholars Pr; 2004.

32. Dg K. Beyond needs-based health funding : resource allocation and equity at the state and area health service levels in New South Wales, Australia. [Internet]. 2009 [cited 2012 Apr 16]. Available from: http://ukpmc.ac.uk/theses/ETH/104

33. Rice N, Smith PC. Capitation and risk adjustment in health care financing: an international progress report. Milbank Q. 2001;79(1):81–113, IV.

34. Kapur D, Webb R, Development UNC on T and, Development HUC for I. Governance-related Conditionalities of the International Financial Institutions. UN; 2000.

35. Smith DH, Gravelle H. The practice of discounting in economic evaluations of healthcare interventions. Int J Technol Assess Health Care. 2001;17(2):236–43.

36. Diderichsen F. Resource Allocation for Health Equity : Issues and Methods. 2004;(September).

37. Rice N, Dixon P, Lloyd DC, Roberts D. Derivation of a needs based capitation formula for allocating prescribing budgets to health authorities and primary care groups in England: regression analysis. BMJ. 2000 Jan 29;320(7230):284–8.

38. Peacock S, Segal L. Capitation funding in Australia: imperatives and impediments. Health Care Manag Sci. 2000 Feb;3(2):77–88.

39. Sheldon TA, Smith GD, Bevan G. Weighting in the dark: resource allocation in the new NHS. BMJ. 1993 Mar 27;306(6881):835–9.

40. Organization WH. World Health Statistics 2010. World Health Organization; 2010.

41. Marmot M. Social determinants of health inequalities. The Lancet. 2005;365(9464):1099–104.

42. Waldram JB, Herring A, Young TK. Aboriginal health in Canada: Historical, cultural, and epidemiological perspectives. Univ of Toronto Pr; 2006.

43. Canada; G of CAA and ND. Royal Commission Report on Aboriginal Peoples [Internet]. 2011 [cited 2012 Apr 16]. Available from: http://www.aadnc-aandc.gc.ca/eng/1307458586498

44. Yusuf S, Reddy S, Ôunpuu S, Anand S. Global burden of cardiovascular diseases. Circulation. 2001;104(23):2855–64.

45. Wing S, Hayes C, Heiss G, John E, Knowles M, Riggan W, et al. Geographic variation in the onset of decline of ischemic heart disease mortality in the United States. American journal of public health. 1986;76(12):1404–8.

46. Black MM, Krishnakumar A. Predicting Longitudinal Growth Curves of Height and Weight Using Ecological Factors for Children with and Without Early Growth Deficiency. J. Nutr. 1999 Feb 1;129(2):539S–543S.

47. Vaidya VG, Alexandro Jr. FJ. Evaluation of some mathematical models for tumor growth. International Journal of Bio-Medical Computing. 1982 Jan;13(1):19–35.

48. Bajzer Z, Marušić M, Vuk-Pavlović S. Conceptual frameworks for mathematical modeling of tumor growth dynamics. Mathematical and Computer Modelling. 1996 Mar;23(6):31–46.

49. Carr-Hill RA, Sheldon TA, Smith P, Martin S, Peacock S, Hardman G. Allocating resources to health authorities: development of method for small area analysis of use of inpatient services. BMJ. 1994 Oct 22;309(6961):1046–9.

50. Eyles J, Birch S, al et. A Needs-Based Methodology for Allocating Health Care Resources in Ontario, Canada: Development and an Application. Social Science & Medicine. 1991;33(4):489.

51. Asada Y. A framework for measuring health inequity. Journal of epidemiology and community health. 2005;59(8):700–5.

52. O'Donnell O, Van Doorsslaer E, Wagstaff A, Lindelöw M. Analyzing health equity using household survey data: a guide to techniques and their implementation [Internet]. World Bank Publications; 2008 [cited 2012 Jul 10]. Available from: http://books.google.ca/books?hl=en&lr=&id=8krsjfKv2vgC&oi=fnd&pg=PR3&dq =A+guide+to+techniques+and+their+implementation+&ots=Ru150T54gU&sig=Ctt rsRwDHiNZ76kyzPY9Opp8S-4

53. Monrad Aas IH. Incentives and financing methods. Health Policy. 1995;34(3):205–20.

54. Durbin J, Cochrane J, Goering P, Macfarlane D. Needs-based planning: evaluation of a level-of-care planning model. J Behav Health Serv Res. 2001 Feb;28(1):67–80.

55. Reid RJ, Roos NP, MacWilliam L, Frohlich N, Black C. Assessing Population Health Care Need Using a Claims-based ACG Morbidity Measure: A Validation Analysis in the Province of Manitoba. Health Services Research. 2002;37(5):1345–64.

56. Fries JF. Aging, natural death, and the compression of morbidity. New England journal of medicine. 1980;303(3):130–5.

57. Fries JF, Baltes PB, BALTES M. Medical perspectives upon successful aging. Successful aging: Perspectives from the behavioral sciences. 1990;35–49.

58. Soldo BJ, Manton KG. Changes in the health status and service needs of the oldest old: Current patterns and future trends. The Milbank Memorial Fund Quarterly. Health and Society. 1985;286–319.

59. Vita AJ, Terry RB, Hubert HB, Fries JF. Aging, health risks, and cumulative disability. New England Journal of Medicine. 1998;338(15):1035–41.

60. Marshall RJ. Mapping disease and mortality rates using empirical Bayes estimators. Applied Statistics. 1991;283–94.

61. Omran AR. The epidemiologic transition: a theory of the epidemiology of population change. The Milbank Memorial Fund Quarterly. 1971;49(4):509–38.

62. Mackenbach JP. The epidemiologic transition theory. J Epidemiol Community Health. 1994 Aug;48(4):329–31.

63. Salomon JA, Murray CJ. The epidemiologic transition revisited: compositional models for causes of death by age and sex. Population and Development Review. 2002;28(2):205–28.

64. Hubert HB, Bloch DA, Oehlert JW, Fries JF. Lifestyle habits and compression of morbidity. The Journals of Gerontology Series A: Biological Sciences and Medical Sciences. 2002;57(6):M347–M351.

65. Duncan TE, Duncan SC, Strycker LA. An introduction to latent variable growth curve modeling: Concepts, issues, and applications [Internet]. Lawrence Erlbaum; 2006 [cited 2012 Jul 23]. Available from: http://books.google.ca/books?hl=en&lr=&id=fAtIoR7aFyMC&oi=fnd&pg=PP1&dq=Growth+curve+modeling&ots=8ffGFSh605&sig=1jzI4zU7Wrbpu4zcnHAJfzZfFGE

66. Duncan TE, Duncan SC. An introduction to latent growth curve modeling. Behavior therapy. 2004;35(2):333–63.

67. Preece MA, Baines MJ. A new family of mathematical models describing the human growth curve. Annals of Human Biology. 1978;5(1):1–24.

68. Burchinal MR, Bailey DB, Snyder P. Using growth curve analysis to evaluate child change in longitudinal investigations. Journal of Early Intervention. 1994;18(4):403–23.

69. Bryk AS, Raudenbush SW. Application of hierarchical linear models to assessing change. Psychological Bulletin. 1987;101(1):147.

70. Zwietering MH, Jongenburger I, Rombouts FM, Van't Riet K. Modeling of the bacterial growth curve. Applied and Environmental Microbiology. 1990;56(6):1875–81.

71. Baranyi J, Roberts TA. A dynamic approach to predicting bacterial growth in food. International journal of food microbiology. 1994;23(3):277–94.

72. Zwietering MH, De Koos JT, Hasenack BE, De Witt JC, Van't Riet K. Modeling of bacterial growth as a function of temperature. Applied and Environmental Microbiology. 1991;57(4):1094–101.

73. Guiot C, Degiorgis PG, Delsanto PP, Gabriele P, Deisboeck TS. Does tumor growth follow a "universal law"? Journal of theoretical biology. 2003;225(2):147–51.

74. Heitjan DF, Manni A, Santen RJ. Statistical analysis of in vivo tumor growth experiments. Cancer Research. 1993;53(24):6042.

75. Laurent S. Rural Canada: Access to health care [Internet]. Parliamentary Research Branch; 2002 [cited 2012 Jul 12]. Available from: http://publications.gc.ca/collections/Collection-R/LoPBdP/PRB-e/PRB0245-e.pdf

76. Wilson NW, Couper ID, De Vries E, Reid S, Fish T, Marais BJ, et al. A critical review of interventions to redress the inequitable distribution of healthcare professionals to rural and remote areas. Rural Remote Health. 2009;9(2):1060.

77. Nemet GF, Bailey AJ. Distance and health care utilization among the rural elderly. Social Science & Medicine. 2000;50(9):1197–208.

78. Fries JF. The compression of morbidity. The Milbank Memorial Fund Quarterly. Health and Society. 1983;397–419.

Distribution of per capita length of stay, by age, across all census divisions


excludes outside values
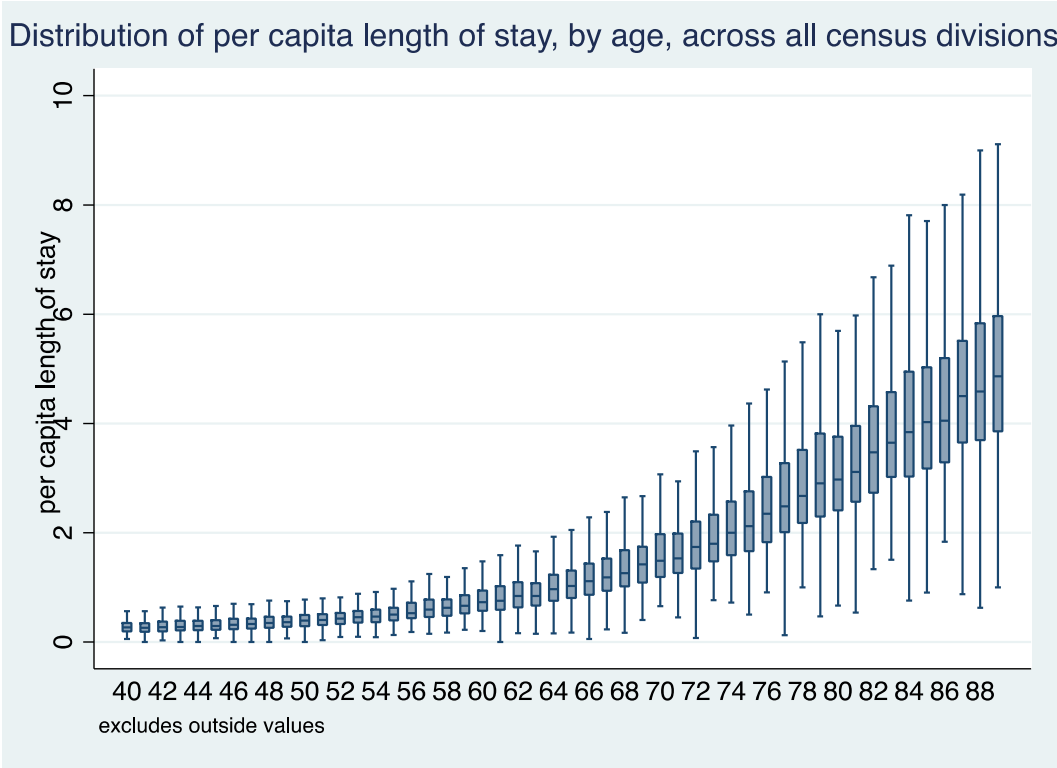
Figure A-1: Box Plot of the Distribution of the Percentiles of the Per Capita Use by Age, for all CDs

Note: Values beyond the 25th and 75th percentiles are excluded. This is for both sexes.

Distribution of per capita length of stay, by age
Females

excludes outside values
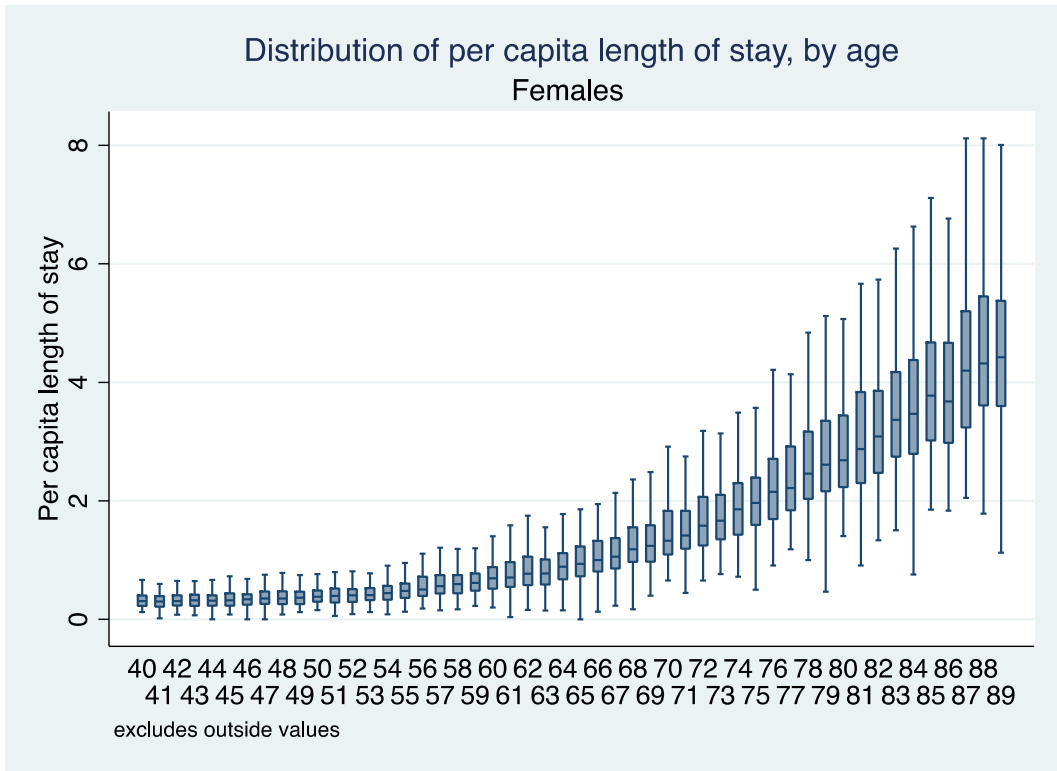
Figure A-2: Box Plot of the Distribution of the Percentiles of Per Capita Use by Age for Females for all CDs

Note: Values beyond the 25th and 75th percentiles are excluded.
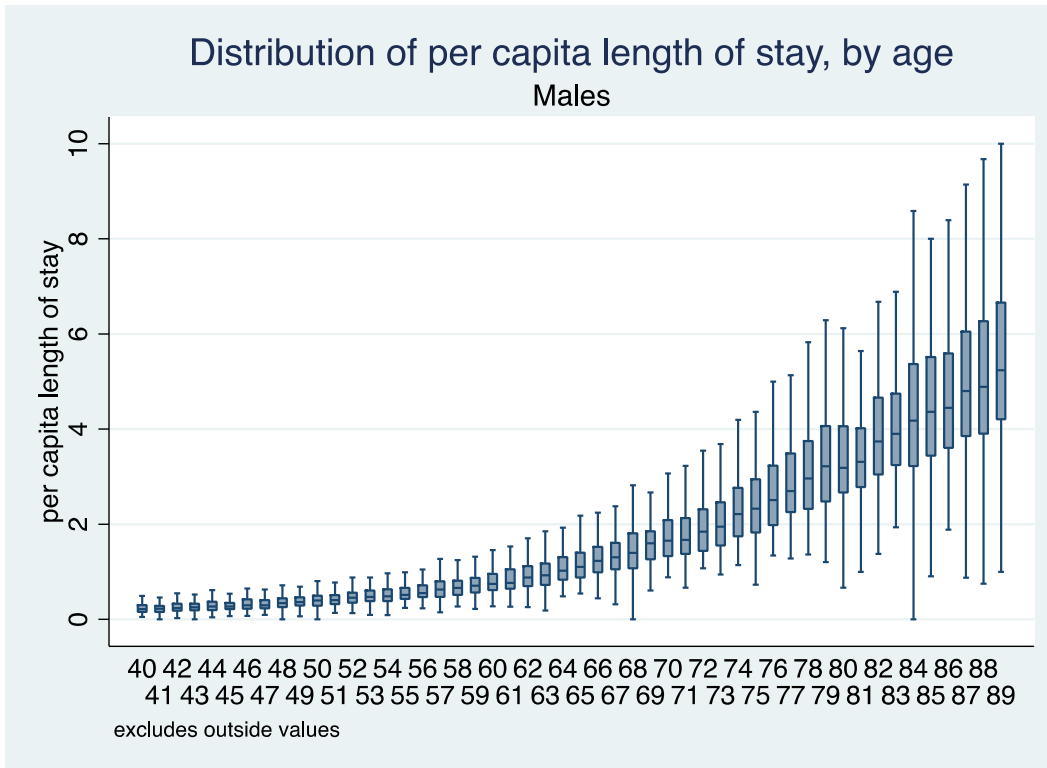
Figure A-3: Box Plot of the Distribution of the Percentiles of Per Capita Use by Age for Males for all CDs

Note: Values beyond the 25th and 75th percentiles are excluded.
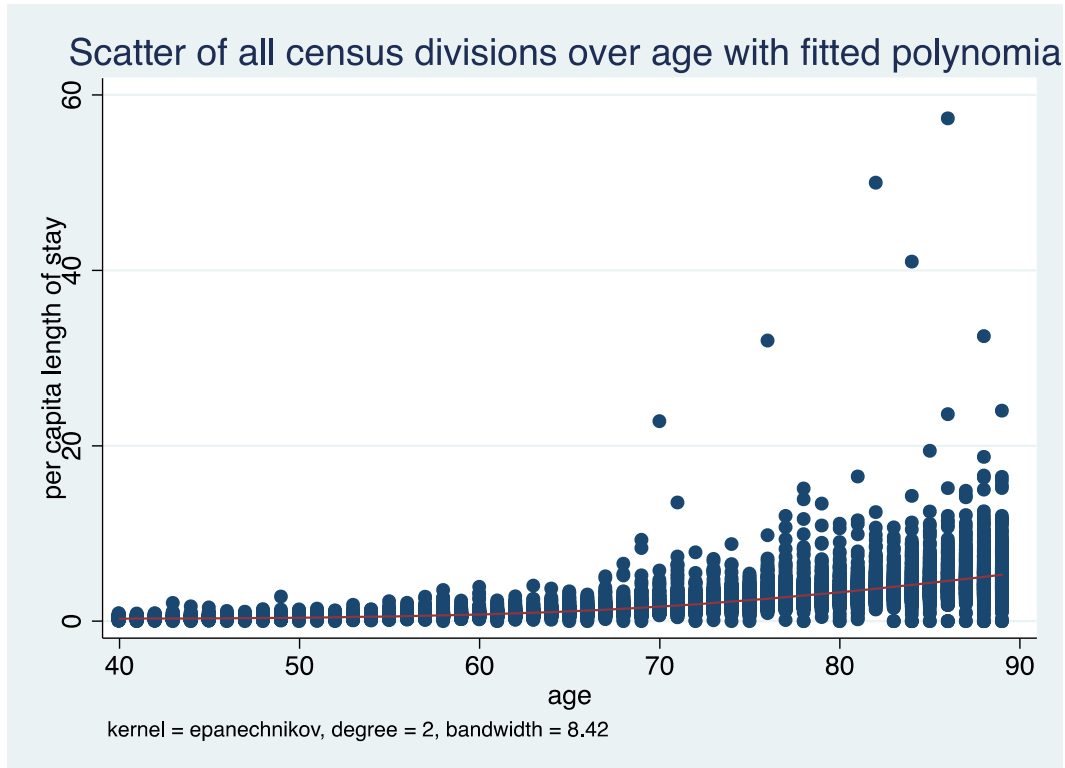
Figure A-4: A Scatter Plot of all Per Capita LOS across all CDs, With a Fitted Non-Linear Regression

Note: The per capita use, by year of age, is fitted with a two-parameter polynomial. All census divisions are included.

Table A-1: The Census Divisions in the Bottom 10th Percentile of Population Size

1011	1213	1216	1218	4604	4606	4608
4610	4616	4619	4620	4623	4704	4804
4818	5943	5945	5957	5959		

Note: The overall population size was used to identify the smallest census divisions, and the Gini analysis was used to confirm that, by excluding these census divisions, the overall age-trajectories of the rate of inpatient hospital use became more uniform and smoothed
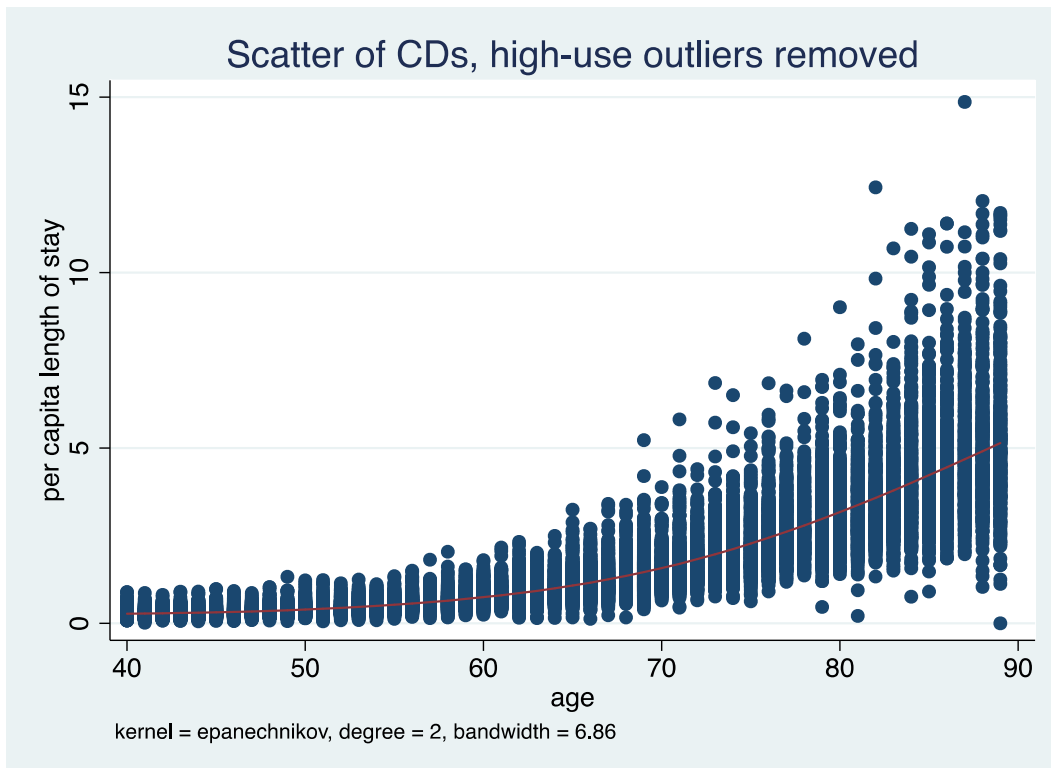
Figure A-5: A Scatter Plot of the Per Capita LOS across CDs, With the Smallest 10th Percentile of Census Divisions Removed

Note: The data are fitted with a two-parameter polynomial, and the excluded CDs are listed in table 1 of Appendix A.
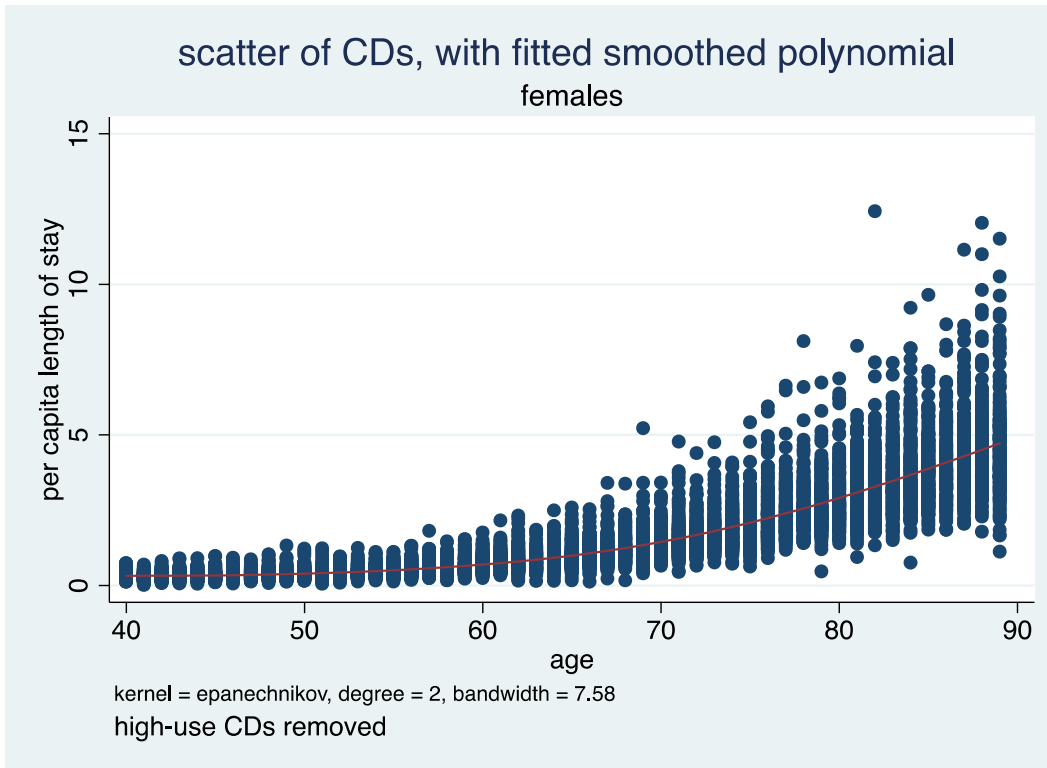
Figure A-6: A Scatter Plot of Per Capita LOS for Females across CDs

Note: The data are fitted with a two-parameter polynomial, and the excluded CDs are listed in table 1 of Appendix A.
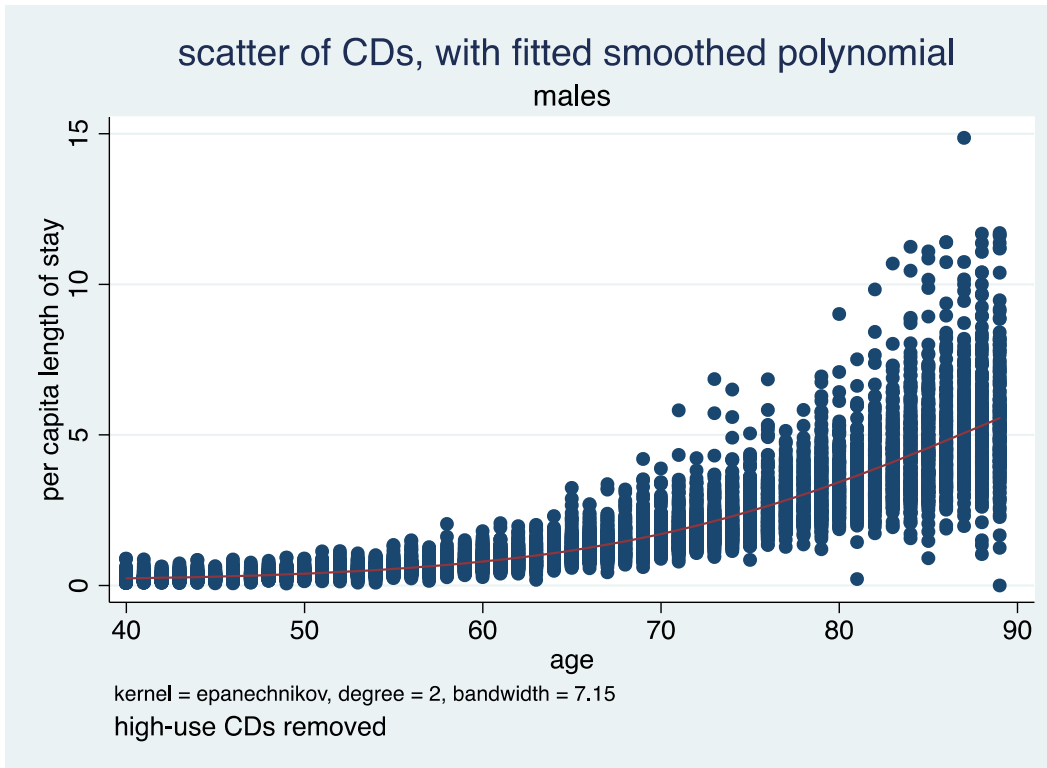
**Figure A-7: A Scatter Plot of Per Capita LOS for Males across CDs**

Note: The data are fitted with a two-parameter polynomial, and the excluded CDs are listed in table 1 of Appendix A.
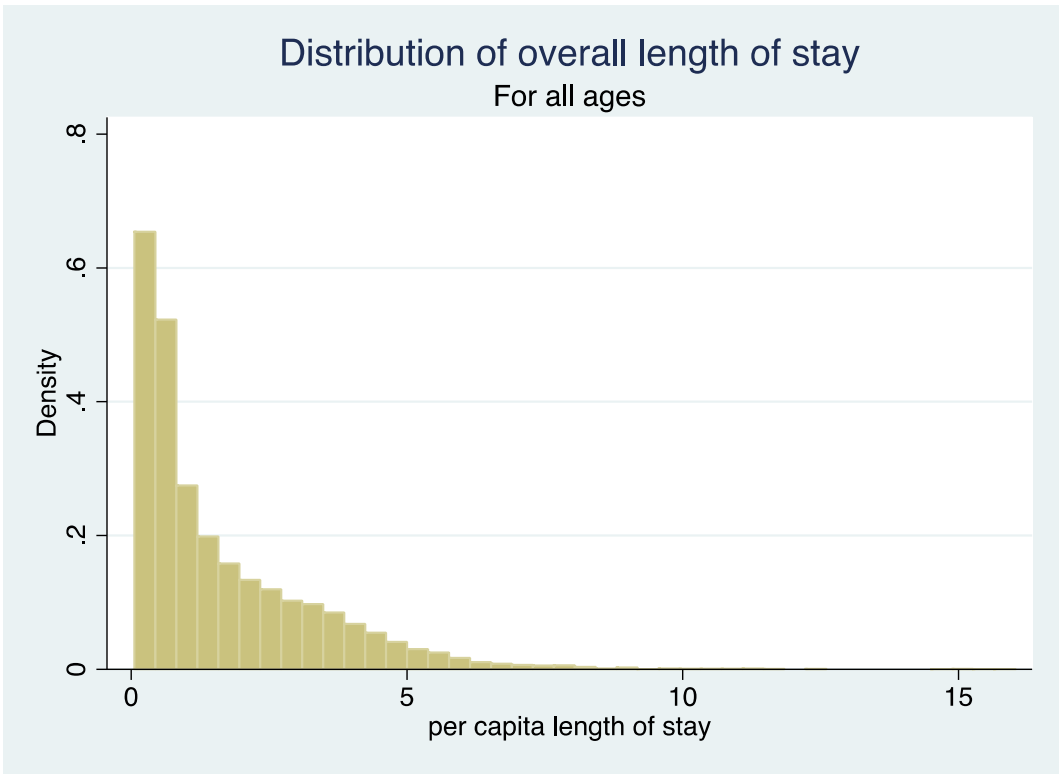
Figure A-8: The Distribution of LOS across CDs for all Ages
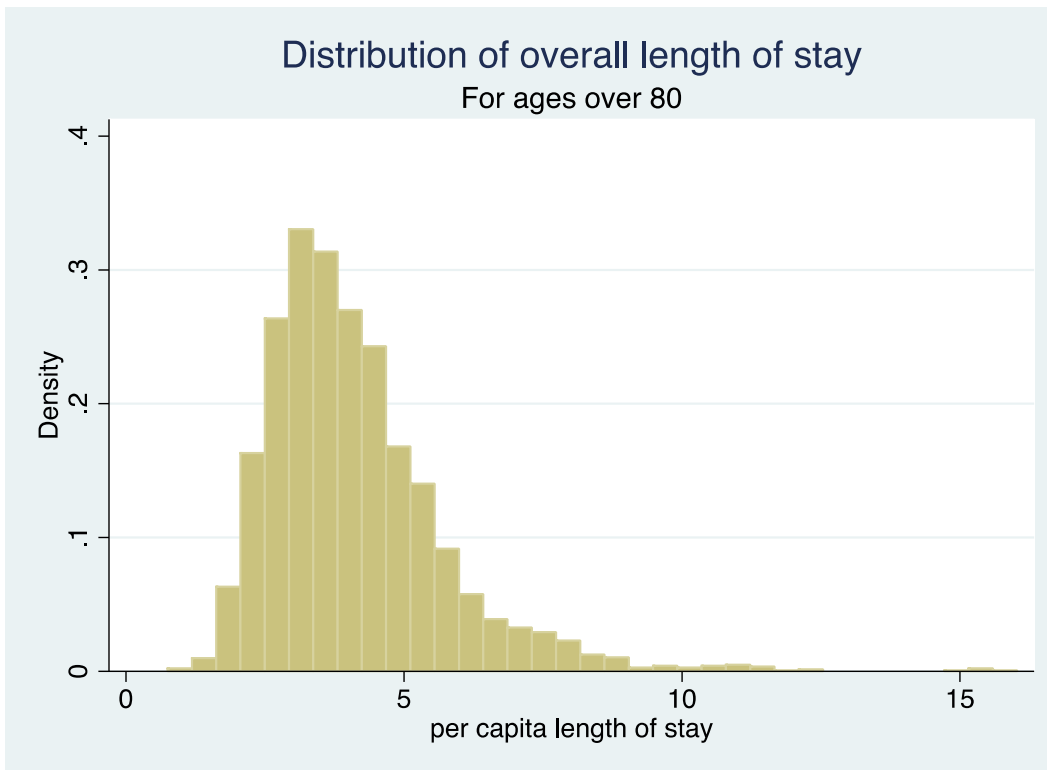
Figure A-9: The Distribution of LOS across CDs for Ages 80 and Above
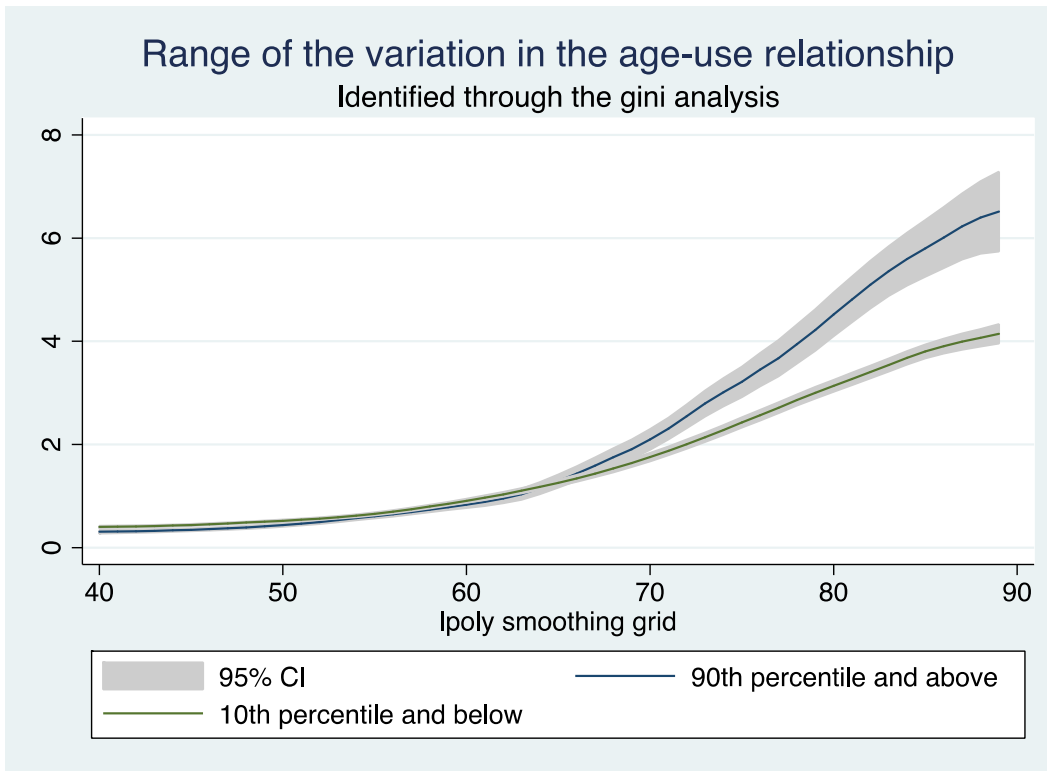
Figure A-10: The Extremes of the Range of Age-Trajectories of the Rate of Inpatient Hospital Use, For Both Sexes

Note: The Gini analysis was used to identify the most extreme range of shapes. The blue line represents CDs with the highest Gini coefficients (that is, the CDs with utilization most concentrated at older ages), and the green line represents those CDs in the bottom 10th percentile of Gini coefficients.
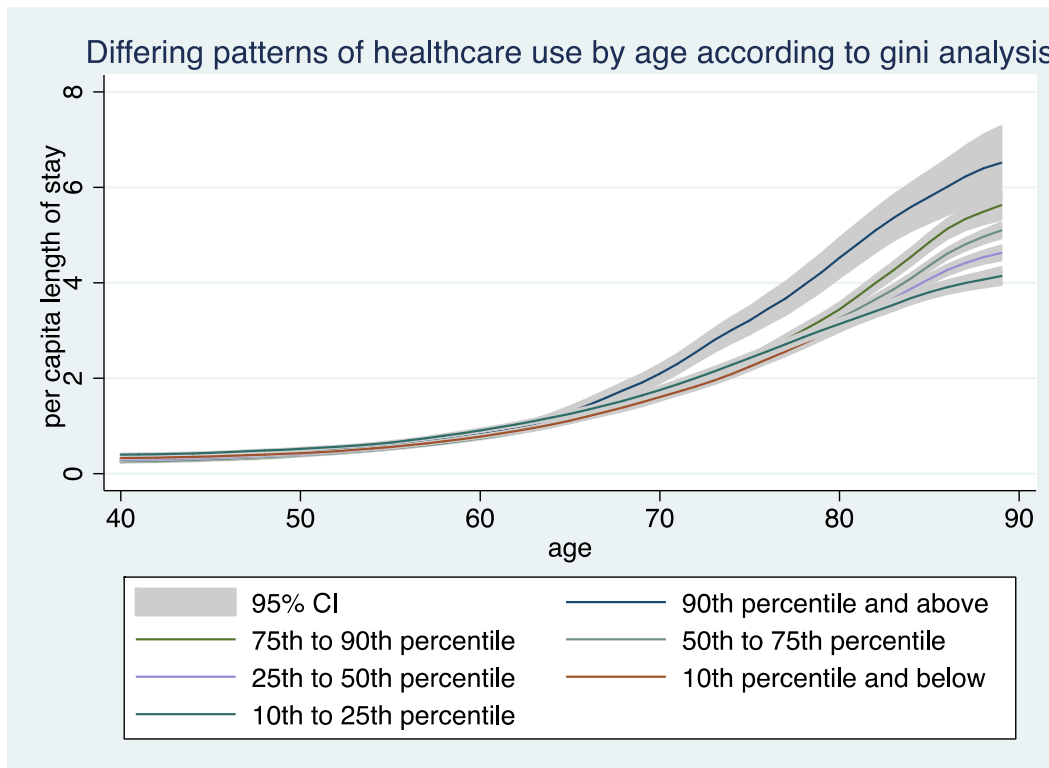
Figure A-11: The Distribution of Age-Trajectories, Identified Through the Gini Analysis

Note: This graph shows the range of shapes, as identified through the Gini analysis. The blue line represents census divisions with the highest Gini coefficients (that is, census divisions with use most concentrated at older ages). The green line represents those census divisions in the bottom tenth percentile of Gini coefficients. The middle three lines all appear to have the same distribution, with slight variation at older ages.
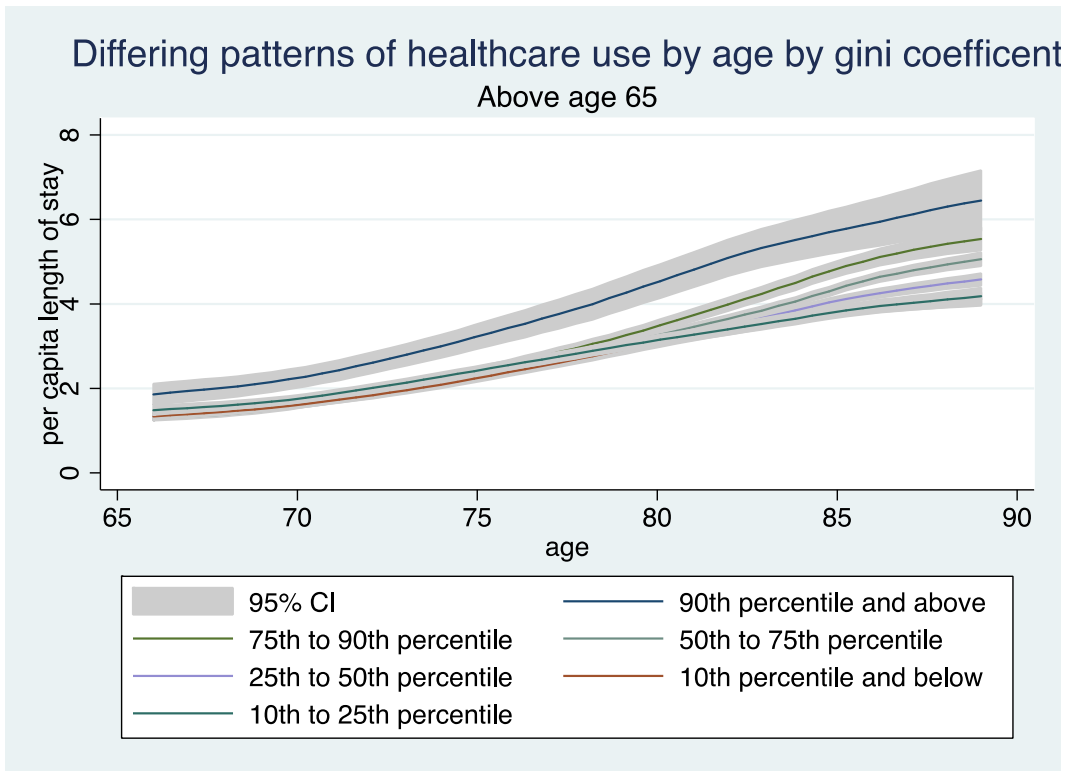
Figure A-12: The Distribution of the Age-Trajectories of the Rate of Inpatient
Hospital Use for Age 65 and Above, As Identified Through the Gini Analysis
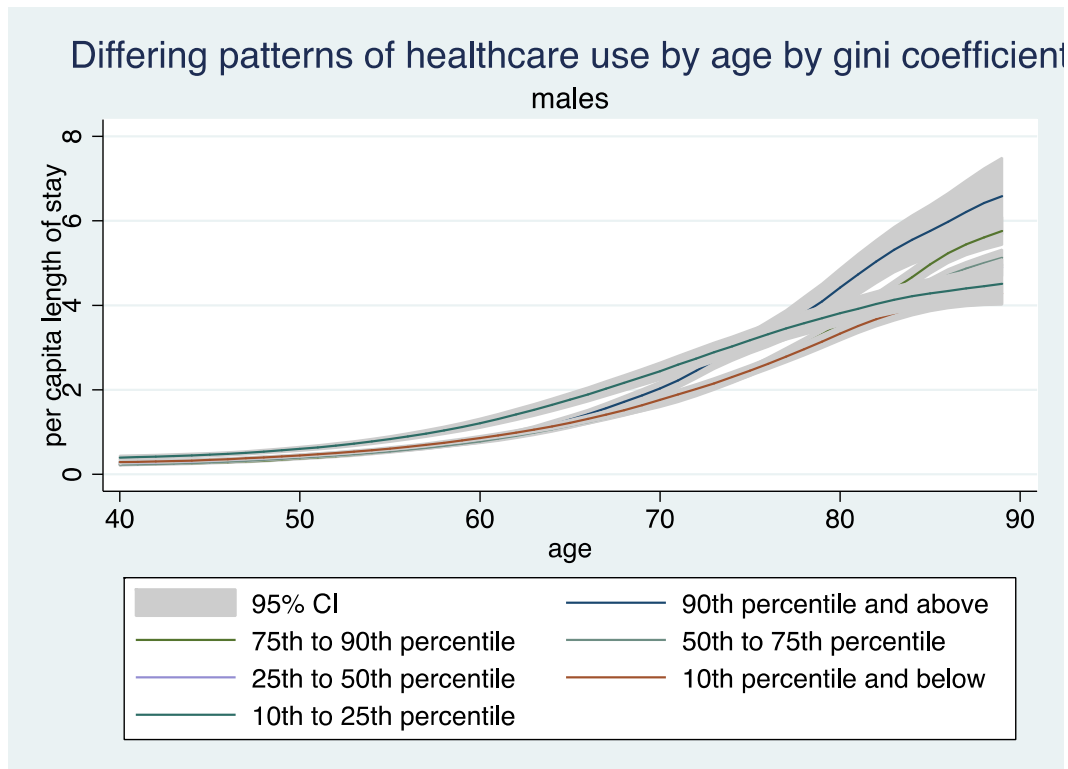
Figure A-13: The Distribution of the Age-Trajectories of the Rate of Inpatient Hospital Use for Males, As Identified Through the Gini Analysis

Note: Figures 14 and 15 serve to show the differences in the shapes of the age-trajectories of the rate of inpatient hospital use between males and females, thus justifying our decision to stratify them in the analysis. The most unusual shape is seen in those between the 10th and 25th percentile. It appears their healthcare use is concentrated at younger ages relative to other CDs. This could be attributable to the fact that some cohorts of men (the bottom socioeconomic groups) are known to have shorter life expectancies. Thus, this may be reflective of that, and the survivor effect of those remaining.
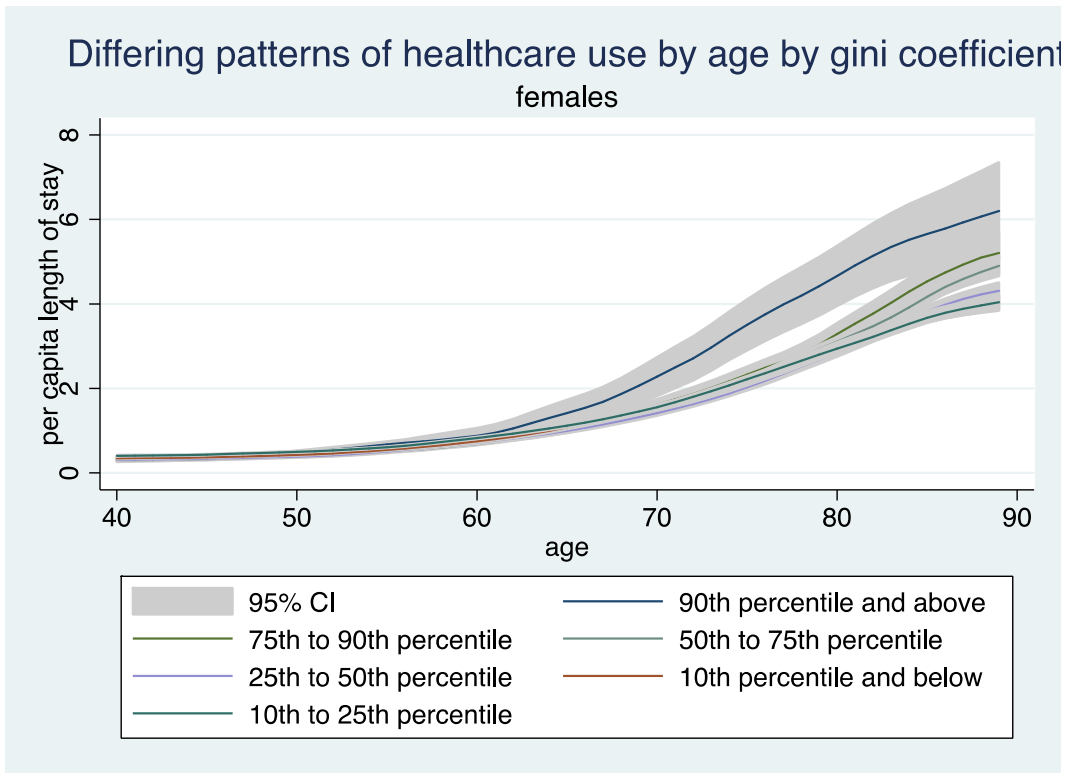
Figure A-14: The Distribution of the Age-Trajectories of the Rate of Inpatient Hospital Use for Females, As Identified Through the Gini Analysis

Note: This graph is much closer in both shape and pattern to the overall graphs. Using the results from Figures 11 to 16, we were able to identify a simple quadratic equation as the appropriate functional form.
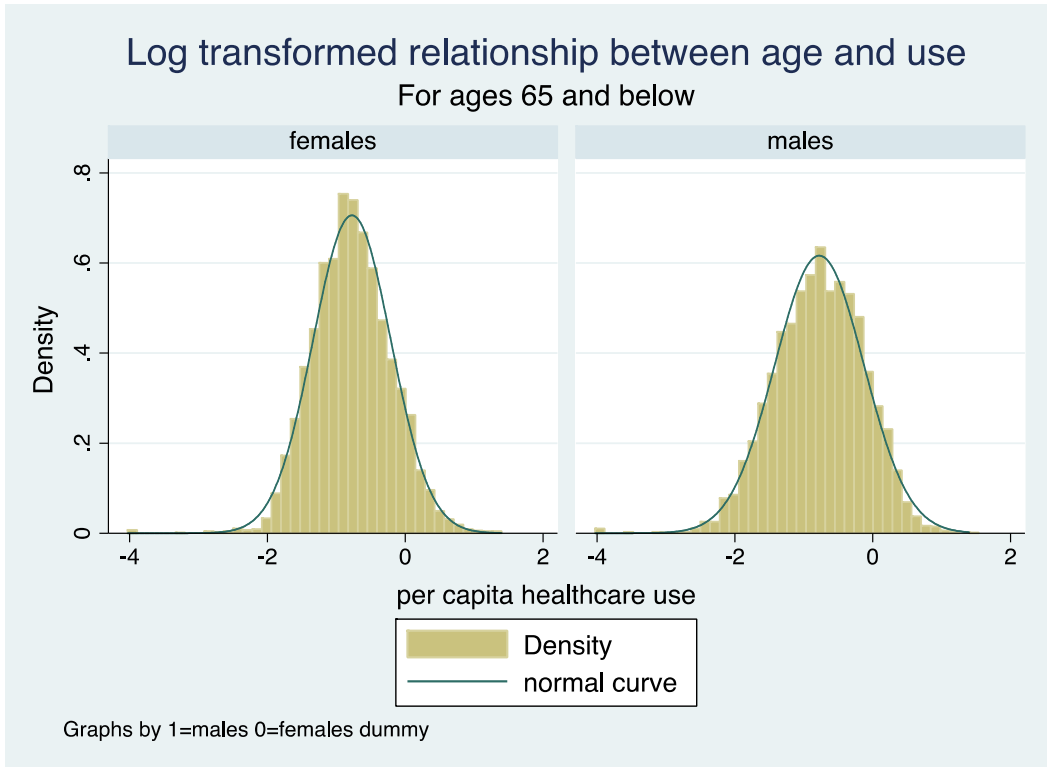
Figure A-15: The Density of the Log-Transformation LOS for Ages 65 and Below, By Sex

Note: Figure 17 presented binomial peaks, one on either side of the origin. Here we can see this is because ages below 65 are responsible for the negative peak.
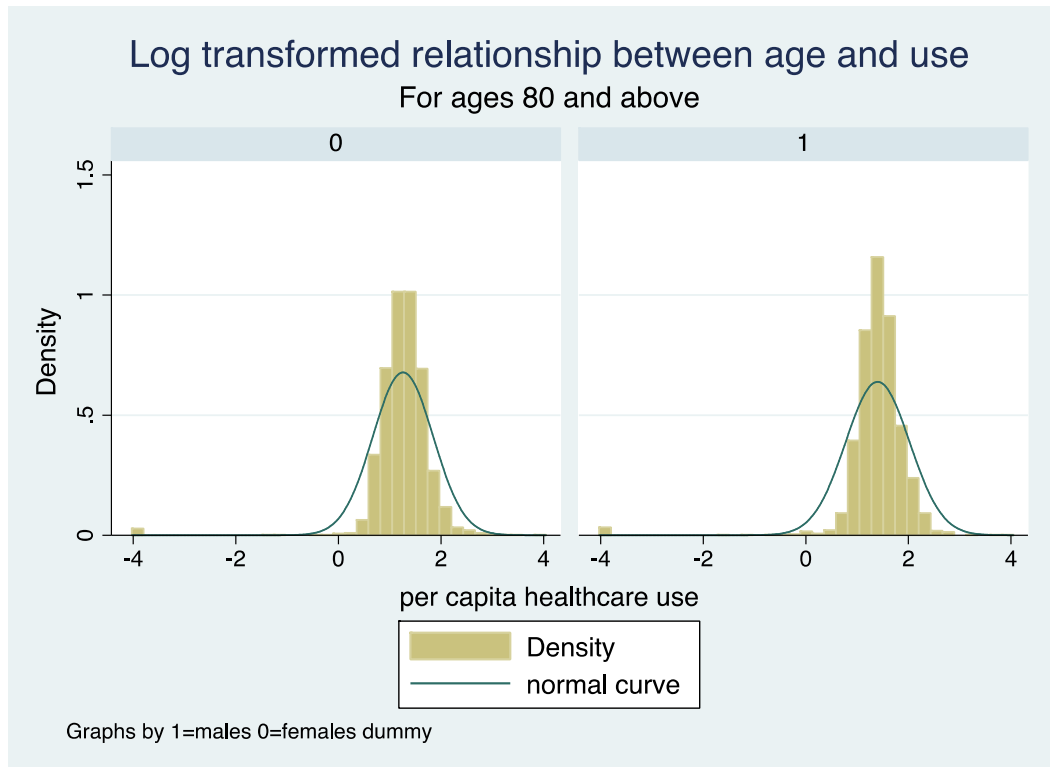
Figure A-16: The Density of the Log-Transformed LOS for Ages 80 and Above, By Sex
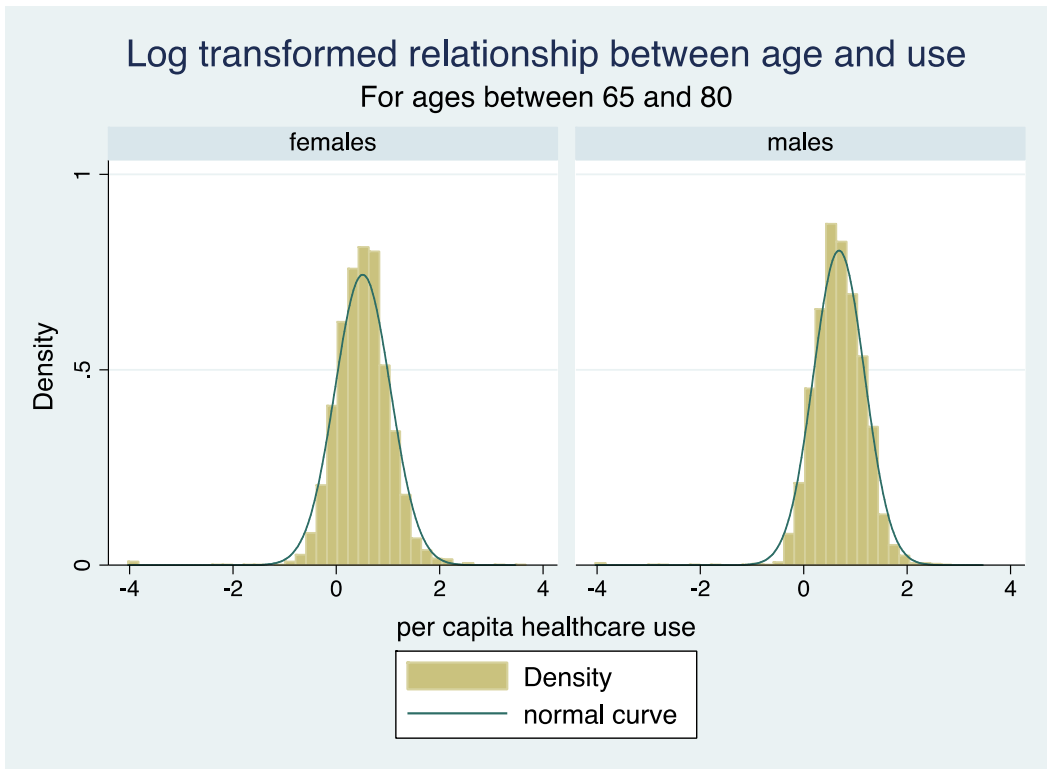
Figure A-17: The Density of the Log-Transformed LOS for Ages between 65 and 80, By Sex

Relationship between log-transformed healthcare use and age

kernel = epanechnikov, degree = 2, bandwidth = 7.43
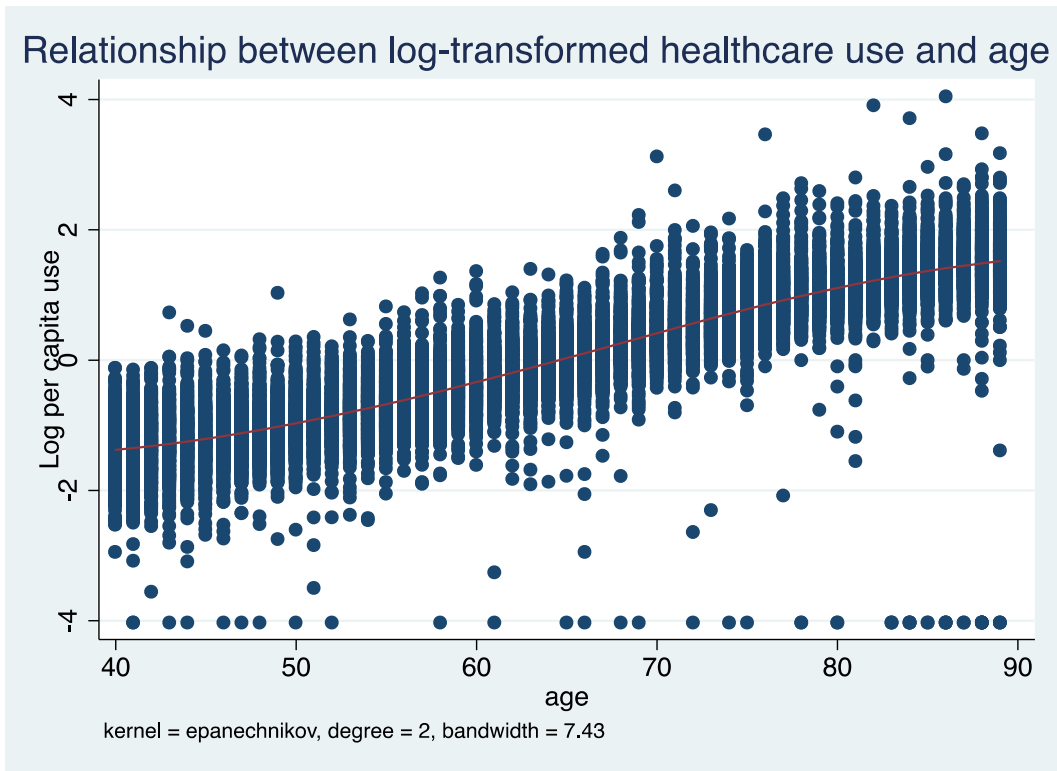
Figure A-18: The Relationship between the Log-Transformed LOS and Age for All CDs
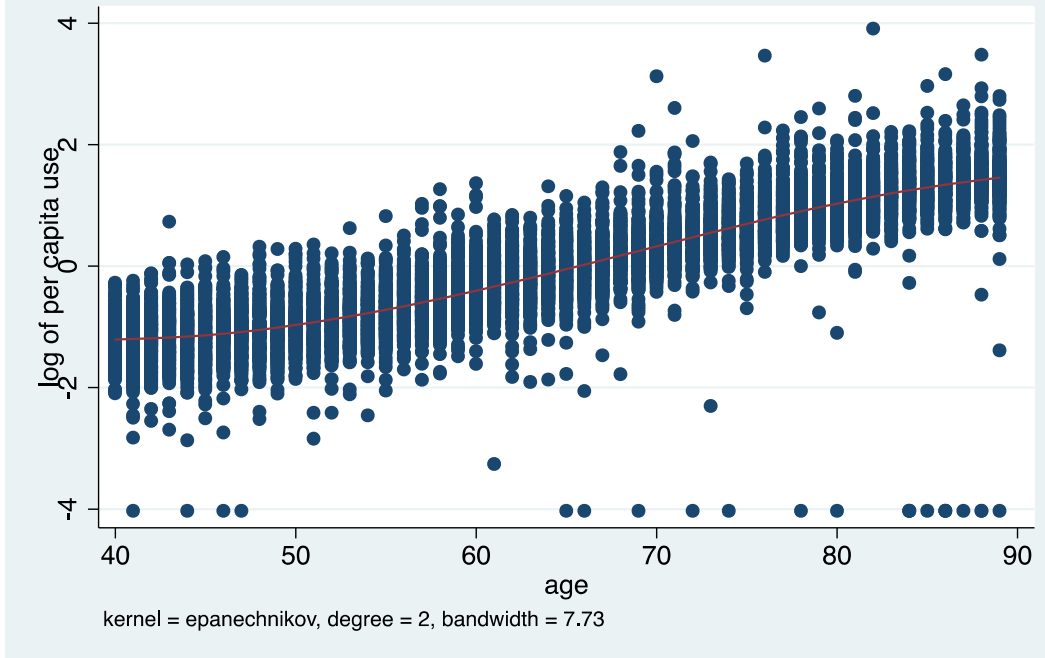
Figure A-19: The Relationship between the Log-Transformed LOS for Females
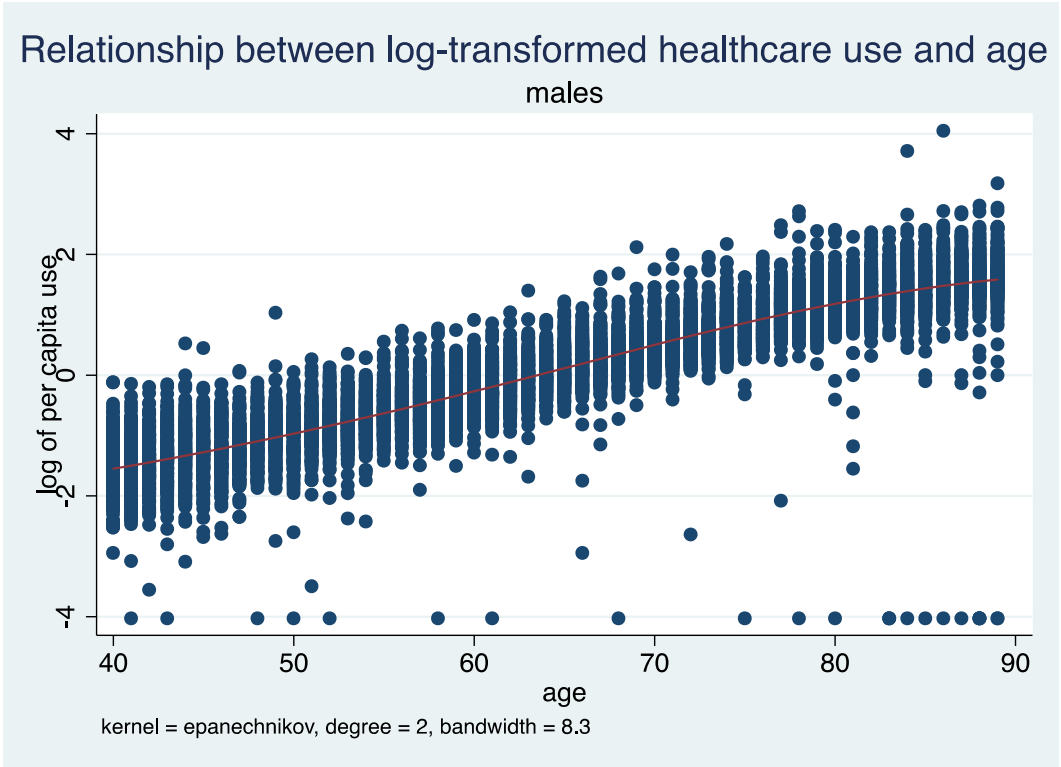
Figure A-20: The Relationship between the Log-Transformed LOS for Males
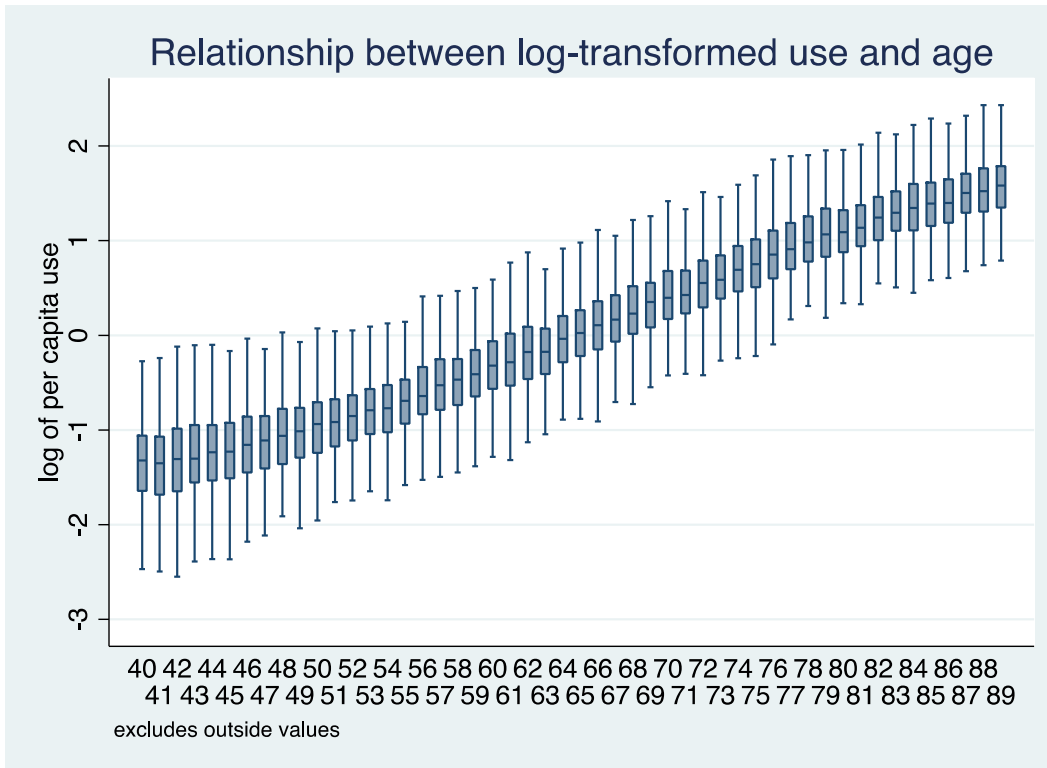
Figure A-21: Box Plot of the Relationship between the Log-Transformed LOS and Age, For Both Sexes

Note: Figures 20 to 23 indicate that the log-transformation linearlizes the relationship between LOS and age.

Relationship between log-transformed use and age
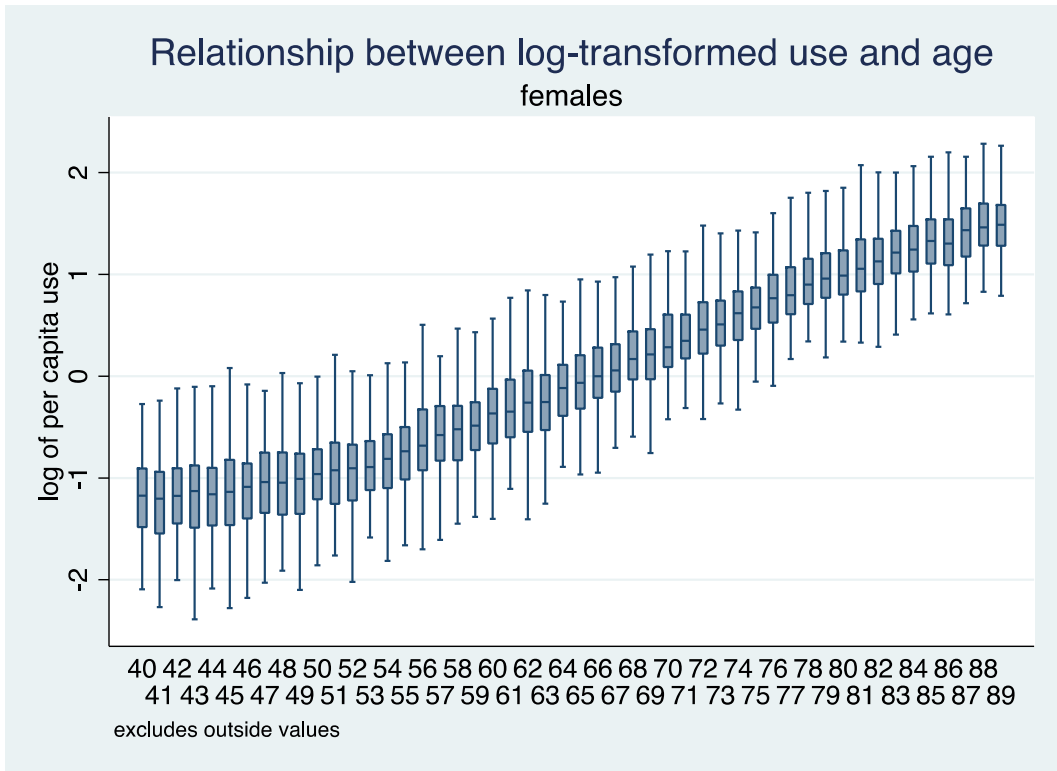
females

Figure A-22: Box Plot of the Relationship between the Log-Transformed LOS and Age for Females
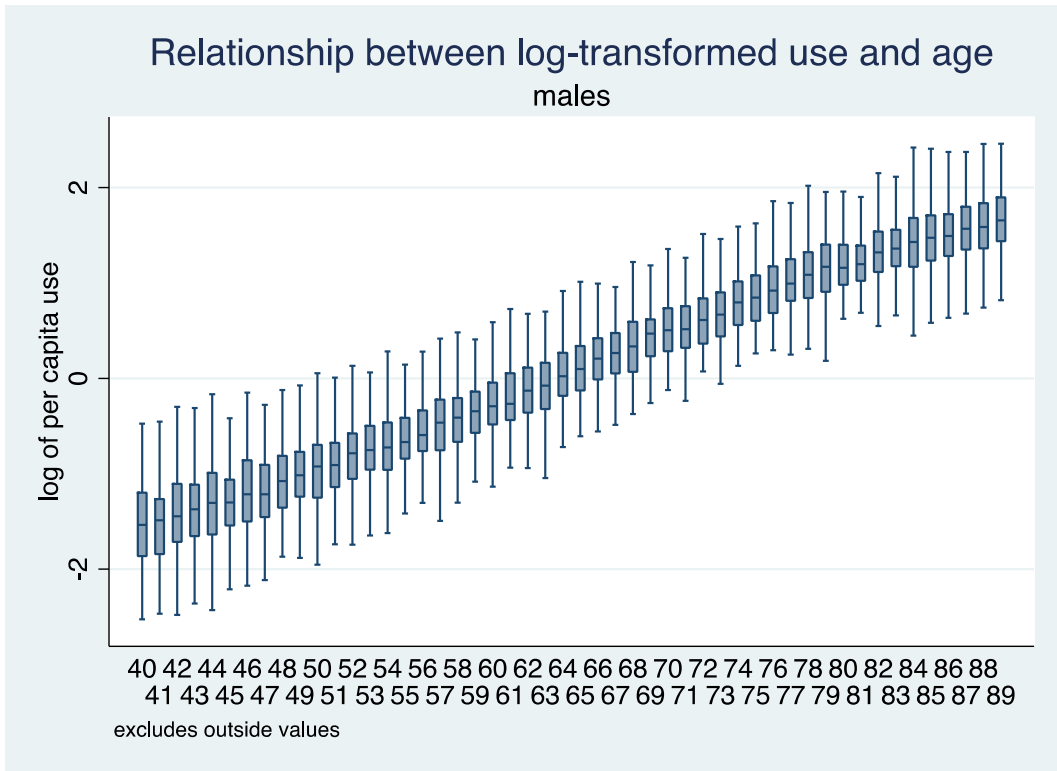
Figure A-23: Box Plot of the Relationship between the Log-Transformed LOS and Age for Males
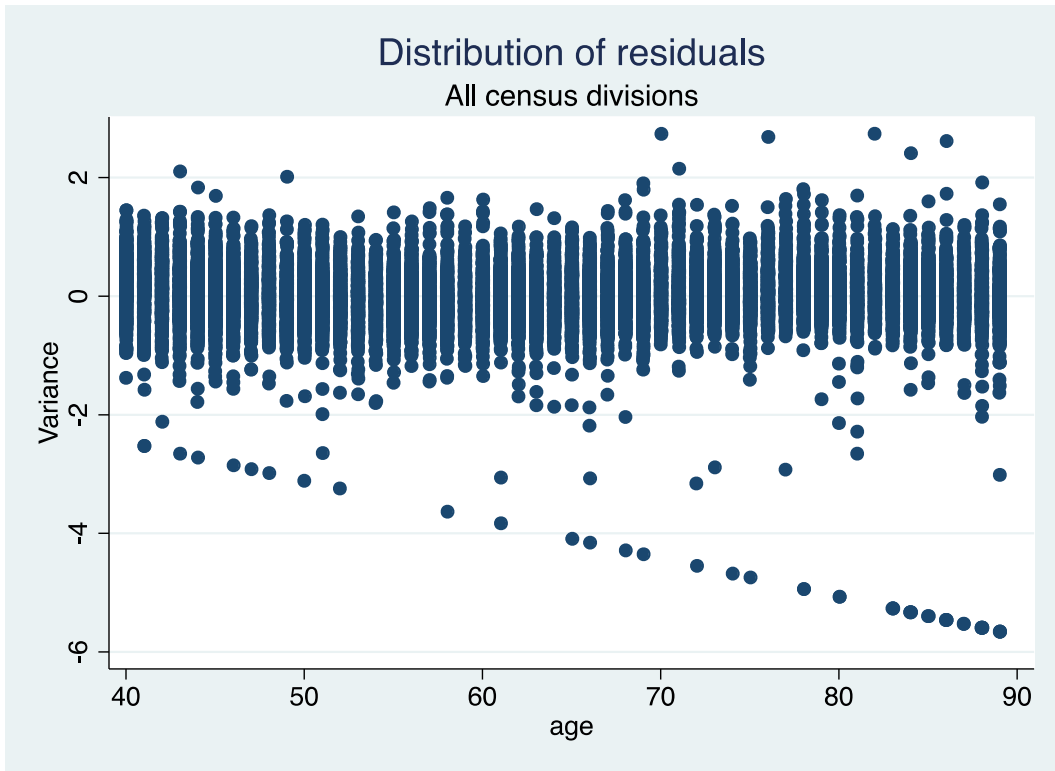
Figure A-24: Distribution of the Residuals from the Predicted Regression Line for All CDs, for Both Sexes

Note: The residuals are evenly concentrated around the origin
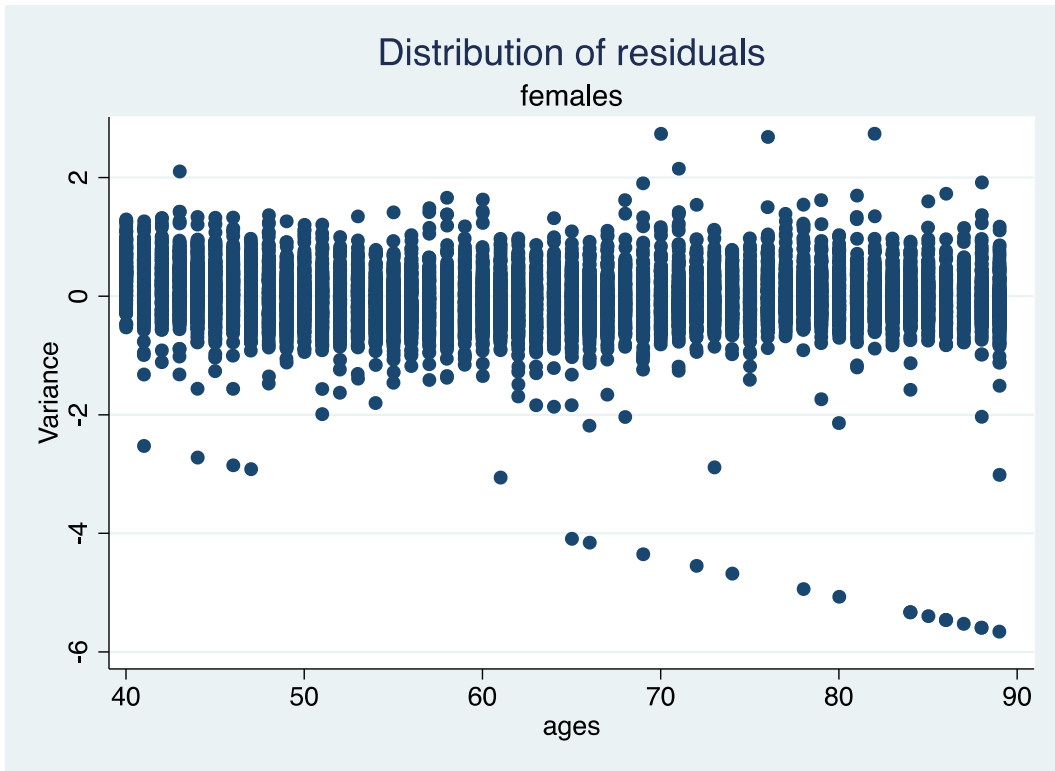
Figure A-25: Distribution of the Residuals from the Predicted Regression Line for All CDs for Females
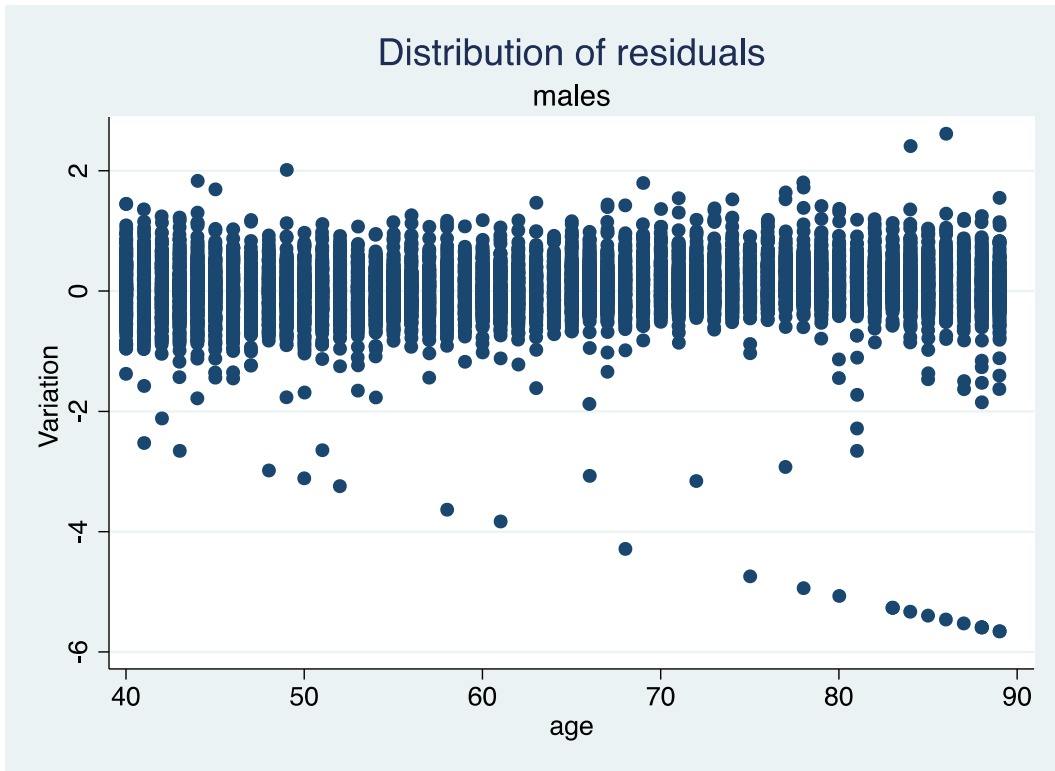
Figure A-26: Distribution of the Residuals from the Predicted Regression Line for All CDs, For Males
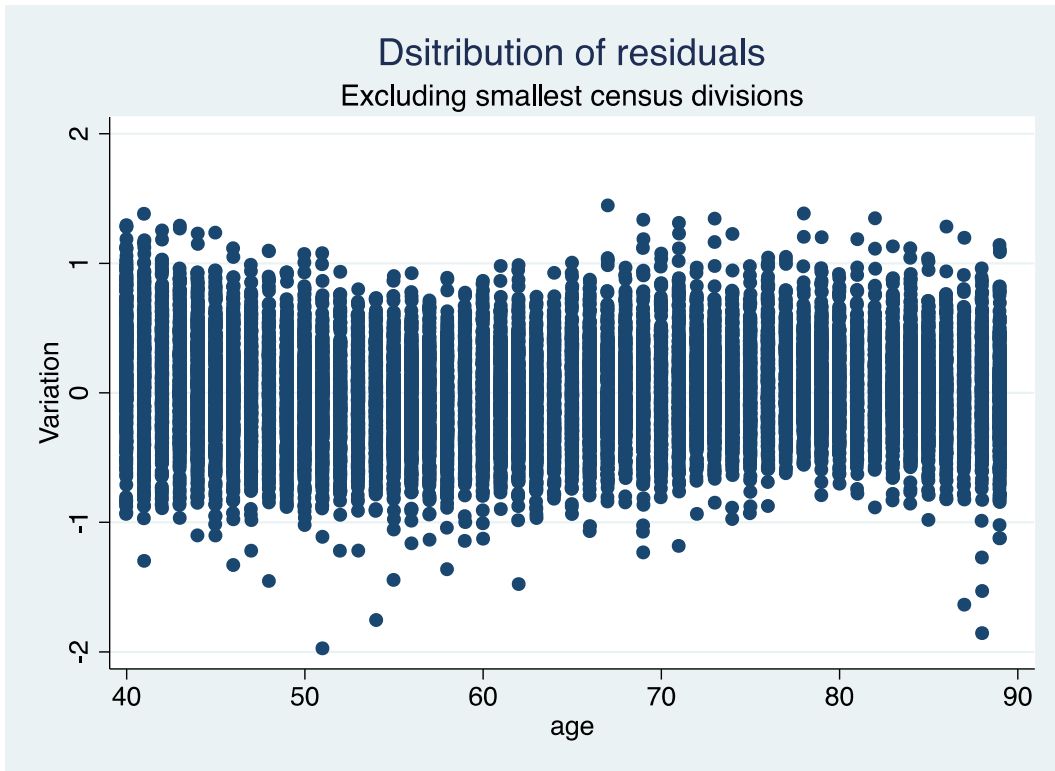
Figure A-27: Distribution of the Residuals from the Predicted Regression Line Excluding the Smallest CDs, For Both Sexes
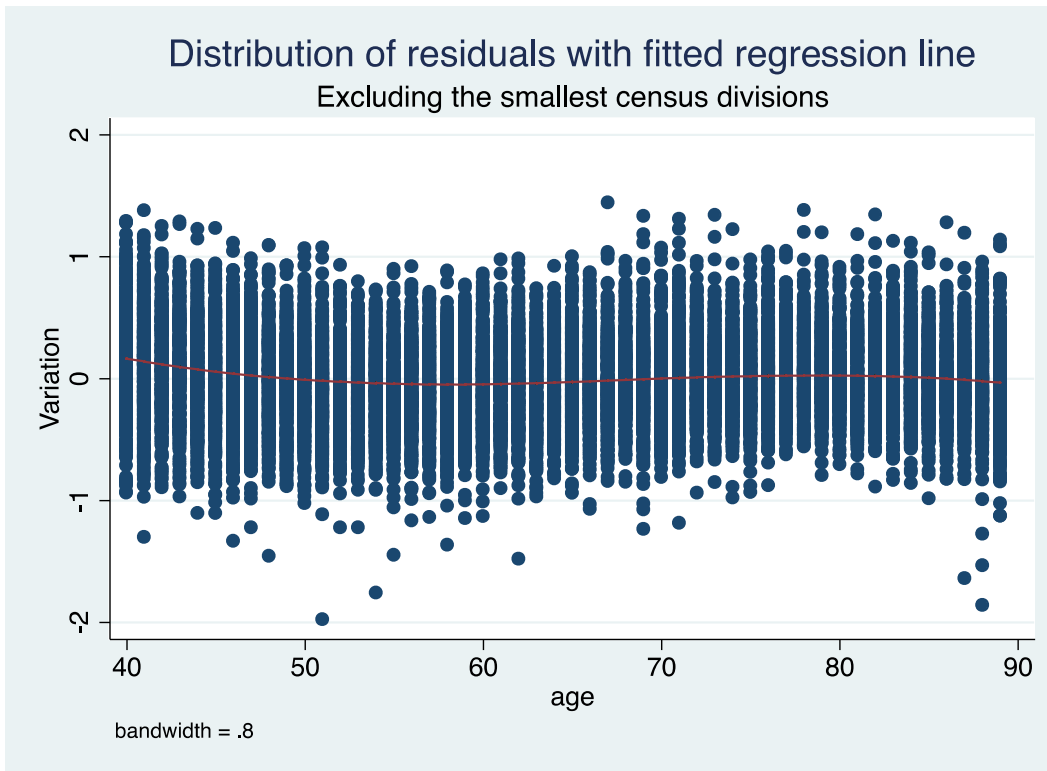
Figure A-28: Distribution of the Residuals from the Predicted Regression Line Excluding the Smallest CDs, With a Fitted Line

Note: The residuals are evenly concentrated around zero, indicating an appropriate fit.
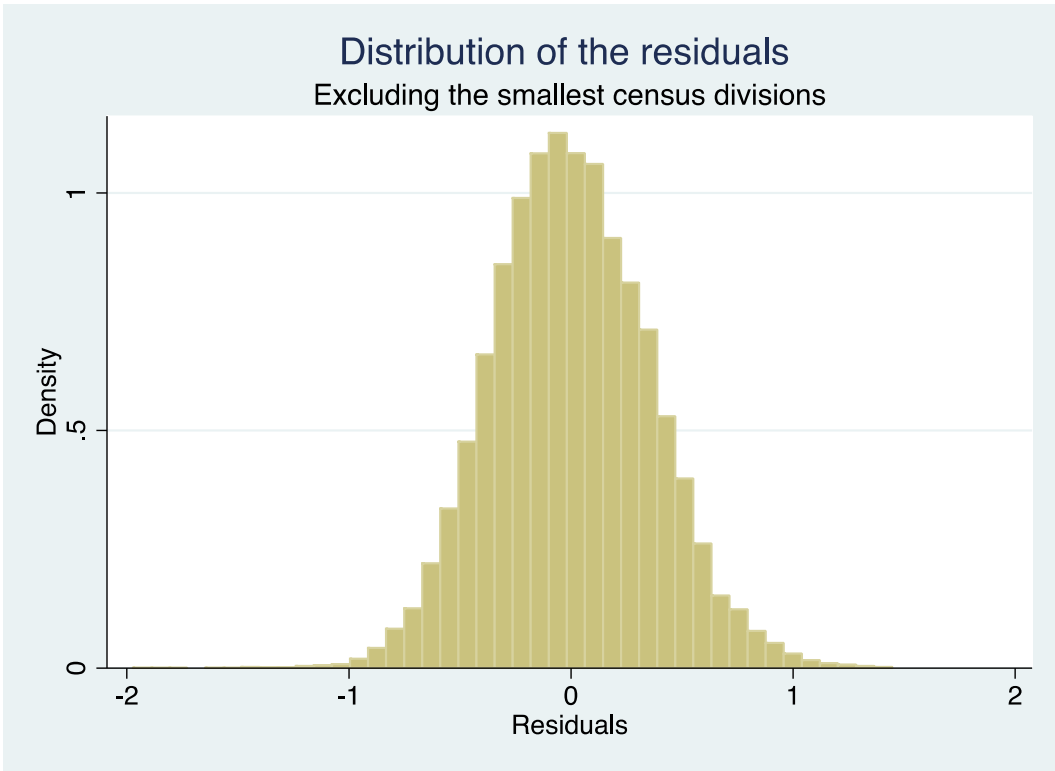
Figure A-29: The Density Distribution of the Residuals

Table A-2: The Results from the Null Model Analysis for Both Sexes Combined

xtreg lnpcLOS age_t, fe i(censusdivision)

Fixed-effects (within) regression	Number of obs	=	17400
Group variable: censusdivi~n	Number of groups	=	174

R-sq: within = 0.9158	Obs per group: min =	100
between = 0.0000	avg =	100.0
overall = 0.8638	max =	100

F(1,17225)   = 187377.27

corr(u_i, Xb) = 0.0000   Prob > F   = 0.0000

```
------------------------------------------------------------------
   lnpcLOS |    Coef.  Std. Err.    t   P>|t|    [95% Conf. Interval]
-------------+----------------------------------------------------
     age_t |  .0658846  .0001522  432.87  0.000   .0655863   .066183
     _cons |  .0574596  .0021978   26.14  0.000   .0531518  .0617675
-------------+----------------------------------------------------
   sigma_u |
.24438211
   sigma_e |  .28972954
       rho |  .41570511   (fraction of variance due to u_i)
------------------------------------------------------------------
```

F test that all u_i=0:   F(173, 17225) =   71.15   Prob > F = 0.0000

Table A-3: The Results from the Random Coefficient Model Analysis for Both Sexes Combined

Performing gradient-based optimization:

Iteration 0:   log likelihood = -2858.3917
Iteration 1:   log likelihood = -2858.3917

Computing standard errors:

Mixed-effects ML regression	Number of obs	=	17400

Group variable: censusdivision Number of groups = 174

Obs per group: min = 100
avg = 100.0
max = 100

Wald chi2(2) = 27894.84
Log likelihood = -2858.3917 Prob > chi2 = 0.0000

```
--------------------------------------------------------------------------
   lnpcLOS |    Coef.  Std. Err.   z   P>|z|   [95% Conf. Interval]
-------------+------------------------------------------------------------
     age_t |  .0660679  .0003992  165.52  0.000   .0652856   .0668502
     agesq |  .0001833  .0000224    8.17  0.000   .0001393   .0002272
     _cons |   .019341  .0197885    0.98  0.328  -.0194438   .0581258
--------------------------------------------------------------------------
```

```
--------------------------------------------------------------------------
  Random-effects Parameters |  Estimate  Std. Err.   [95% Conf. Interval]
-----------------------------+--------------------------------------------
censusdivi~n: Unstructured   |
              sd(age_t) |  .0049075  .0003028   .0043485   .0055385
              sd(agesq) |  .0002564  .0000183   .0002229   .0002949
              sd(_cons) |  .2577624   .01417    .2314337   .2870864
       corr(age_t,agesq) | -.1224253  .0929178  -.2985519   .0617653
       corr(age_t,_cons) | -.2356103  .0781327  -.3818804  -.0778256
       corr(agesq,_cons) | -.3994825  .0743886  -.5345671   -.244494
-----------------------------+--------------------------------------------
            sd(Residual) |  .2744398  .0014937   .2715277   .2773831
--------------------------------------------------------------------------
```

LR test vs. linear regression: chi2(6) = 9614.53 Prob > chi2 = 0.0000

Note: LR test is conservative and provided only for reference.

Note: This model treated the slope and intercept as random effects but did not include any census division-level variables

Table A-4: The Results from the Full Model Analysis for Both Sexes Combined

Iteration 0: log likelihood = -2729.4699	
Iteration 1: log likelihood = -2729.4699	

Computing standard errors:

Mixed-effects ML regression	Number of obs = 17400
Group variable: censusdivision	Number of groups = 174

Obs per group: min =   100
avg =   100.0
max =   100

Wald chi2(24)   = 43319.86
Log likelihood = -2729.4699   Prob > chi2   =   0.0000

```
---------------------------------------------------------------------------------
        lnpcLOS |   Coef.  Std. Err.    z   P>|z|    [95% Conf.
Interval]
--------------------------+------------------------------------------------------
           age_t |  .104759  .0500584   2.09  0.036    .0066464   .2028716
           agesq |  .0001833  .0000224   8.17  0.000    .0001393   .0002272
        divorced | -.0343346  .0105813  -3.24  0.001   -.0550735  -.0135957
   loneparentfams |  .0139595   .00443    3.15  0.002    .005277   .0226421
        regrepair | -.0087213  .0030031  -2.90  0.004   -.0146073  -.0028353
         can_cit | -.0611891  .0162203  -3.77  0.000   -.0929803  -.0293979
 first_generation | -.0307293  .0050035  -6.14  0.000   -.0405361  -.0209226
         ab_iden |  .0015762  .0017051   0.92  0.355   -.0017658   .0049182
 em_participation |  .0274422  .0059687   4.60  0.000    .0157437   .0391406
         em_rate | -.024335  .0049931  -4.87  0.000   -.0341213  -.0145486
           ed_HS |  .0173364   .003897   4.45  0.000    .0096985   .0249744
     vis_minority |  .0076773  .0045444   1.69  0.091   -.0012296   .0165841
       income_fam |  3.35e-06  2.23e-06   1.50  0.133   -1.02e-06   7.71e-06
                 |
  c.divorced#c.age_t | -.0016374  .0003213  -5.10  0.000   -.0022671  -.0010076
                 |
 c.loneparentfams#c.age_t | -.000278  .0001345  -2.07  0.039   -.0005417  -.0000144
```

```
                          |
       c.regrepair#c.age_t |  -.0000405   .0000912    -0.44   0.657    -.0002192    .0001382
                          |
        c.can_cit#c.age_t |  -.0002274   .0004925    -0.46   0.644    -.0011928    .0007379
                          |
  c.first_generation#c.age_t | -9.46e-06   .0001519    -0.06   0.950    -.0003072    .0002883
                          |
         c.ab_iden#c.age_t |  -.0001111   .0000518    -2.15   0.032    -.0002126   -9.63e-06
                          |
  c.em_participation#c.age_t | -.0002369   .0001812    -1.31   0.191    -.0005921    .0001184
                          |
         c.em_rate#c.age_t |   .0001204   .0001516     0.79   0.427    -.0001767    .0004176
                          |
          c.ed_HS#c.age_t |  -.0000798   .0001183    -0.67   0.500    -.0003117    .0001521
                          |
    c.vis_minority#c.age_t |  -7.92e-06    .000138    -0.06   0.954    -.0002784    .0002625
                          |
     c.income_fam#c.age_t |   1.78e-07   6.76e-08     2.64   0.008     4.58e-08    3.11e-07
                          |
                    _cons |   5.827534   1.648514     3.54   0.000     2.596506    9.058562
------------------------------------------------------------------------------------------------
```

```
------------------------------------------------------------------------------
  Random-effects Parameters |  Estimate   Std. Err.    [95% Conf.
Interval]
----------------------------+-------------------------------------------------
-
censusdivi~n: Unstructured  |
              sd(age_t) |   .0027906   .0002554     .0033218    .0043257
              sd(agesq) |   .0002564   .0000183     .0002229    .0002949
              sd(_cons) |   .1336364   .0078673     .1190732    .1499807
        corr(age_t,agesq) |  -.1290497   .1136813    -.3420022    .0965101
        corr(age_t,_cons) |  -.2402891   .0846875    -.3979714   -.0688165
        corr(agesq,_cons) |   .0237012   .1080213    -.1859426    .2312809
----------------------------+-------------------------------------------------
-
           sd(Residual) |   .2744398   .0014937     .2715277    .2773831
------------------------------------------------------------------------------
LR test vs. linear regression:     chi2(6) =  4057.42   Prob > chi2 = 0.0000
```

Note: LR test is conservative and provided only for
reference.

Table A-5: Model 1 Regression for Females

Random-effects GLS regression			Number of obs	=	8700
Group variable: censusdivi~n			Number of groups	=	174

R-sq: within = 0.0000	Obs per group: min =	50
between = 0.0000	avg =	50.0
overall = 0.8539	max =	50

Wald chi2(2)   = 101188.94

corr(u_i, X)   = 0 (assumed)   Prob > chi2   =   0.0000

```
------------------------------------------------------------------
   lnpcLOS |    Coef.  Std. Err.    z   P>|z|   [95% Conf. Interval]
-------------+----------------------------------------------------
-
     age_t |  .0626209  .0001969  318.09  0.000   .0622351   .0630068
     agesq |  .0004103  .0000152   26.97  0.000   .0003805   .0004401
     _cons | -.0706492  .0204395   -3.46  0.001  -.1107098  -.0305886
-------------+----------------------------------------------------
-
   sigma_u |
.26372691
   sigma_e |  .26418995
       rho |  .4991229   (fraction of variance due to u_i)
------------------------------------------------------------------
```

Table A-6: Model 1 Regression for Males

Random-effects GLS regression			Number of obs	=	8700
Group variable: censusdivi~n			Number of groups	=	174

R-sq: within = 0.0000	Obs per group: min =	50
between = 0.0000	avg =	50.0
overall = 0.8848	max =	50

Wald chi2(2)   = 107973.50

corr(u_i, X)   = 0 (assumed)   Prob > chi2   =   0.0000

```
------------------------------------------------------------------
   lnpcLOS |    Coef.  Std. Err.    z   P>|z|   [95% Conf. Interval]
-------------+----------------------------------------------------
-
```

```
        age_t |  .0695149  .0002123  327.39  0.000   .0690987   .069931
        agesq | -.0000438  .0000164   -2.67  0.008  -.0000759 -.0000116
        _cons |  .1093312  .0176063    6.21  0.000   .0748234   .143839
-------------+----------------------------------------------------------
-
     sigma_u |
.22424073
     sigma_e |  .2849435
         rho |  .38245486   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

Table A-7: Model 2 Regression for Females

Iteration 0:   log likelihood = -808.27762
Iteration 1:   log likelihood = -808.27762

Computing standard errors:

Mixed-effects ML regression	Number of obs = 8700
Group variable: censusdivision	Number of groups = 174

Obs per group: min =   50

avg =
50.0

max =
50

Wald chi2(2)   = 22338.36

Log likelihood = -808.27762   Prob > chi2   =   0.0000

```
------------------------------------------------------------------------------
   lnpcLOS |   Coef.  Std. Err.    z   P>|z|    [95% Conf. Interval]
-------------+----------------------------------------------------------
-
     age_t |  .0626209  .0004316  145.10  0.000   .061775  .0634668
     agesq |  .0004103  .0000239   17.16  0.000  .0003634  .0004572
     _cons | -.0706492  .0216876   -3.26  0.001 -.1131561 -.0281423
------------------------------------------------------------------------------


------------------------------------------------------------------------------
```

```
  Random-effects Parameters |  Estimate  Std. Err.   [95% Conf.
Interval]
---------------------------+-----------------------------------------------
-
censusdivi~n: Unstructured   |
              sd(age_t) |  .0051421   .000338     .0045205   .0058491
              sd(agesq) |  .0002527   .0000211    .0002145   .0002977
              sd(_cons) |  .2811789   .0156029    .2522021   .3134849
        corr(age_t,agesq) |  -.202784   .102429    -.3927105   .0037322
        corr(age_t,_cons) |  -.2385651  .0810547    -.3899181  -.0746615
        corr(agesq,_cons) |  -.436727   .0784138    -.5770902  -.2713027
---------------------------+-----------------------------------------------
-
             sd(Residual) |  .2485517   .0019435    .2447716   .2523901
-----------------------------------------------------------------------------
LR test vs. linear regression:     chi2(6) =  5900.70   Prob > chi2 = 0.0000
```

Table A-8: Model 2 Regression for Males

Performing gradient-based optimization:

Iteration 0:   log likelihood = -1434.5945
Iteration 1:   log likelihood = -1434.5945

Computing standard errors:

```
Mixed-effects ML regression              Number of obs    =     8700
Group variable: censusdivision           Number of groups  =      174

                              Obs per group: min =     50
                                             avg =
50.0
                                             max =
50

                                  Wald chi2(2)    =  23623.55
Log likelihood = -1434.5945                Prob > chi2     =   0.0000

-----------------------------------------------------------------------------
   lnpcLOS |    Coef.  Std. Err.    z   P>|z|    [95% Conf. Interval]
-----------+-----------------------------------------------------------------
-
```

age_t	.0695149	.0004523	153.69	0.000	.0686284	.0704014
agesq	-.0000438	.0000279	-1.57	0.117	-.0000985	.0000109
_cons	.1093312	.0186042	5.88	0.000	.0728675	.1457948

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]

----------------------------+--
-

censusdivi~n: Unstructured |

sd(age_t)	.0053542	.0003565	.0046991	.0061007
sd(agesq)	.0003068	.0000237	.0002637	.000357
sd(_cons)	.2387419	.0135229	.2136559	.2667734
corr(age_t,agesq)	-.0244745	.1013977	-.2196937	.1726293
corr(age_t,_cons)	-.2299867	.082789	-.384646	-.0627672
corr(agesq,_cons)	-.3674763	.0800064	-.5130004	-.2014191

----------------------------+--
-

sd(Residual)	.2677995	.002094	.2637267	.2719353

LR test vs. linear regression: chi2(6) = 4147.69   Prob > chi2 = 0.0000

Note: LR test is conservative and provided only for reference.

Table A-9: Model 3 Regression for Females

Performing gradient-based optimization:

Iteration 0:   log likelihood = -697.05461
Iteration 1:   log likelihood = -697.05458

Computing standard errors:

Mixed-effects ML regression	Number of obs = 8700
Group variable: censusdivision	Number of groups = 174

Obs per group: min =   50

avg =
50.0

max =

Wald chi2(23)   = 33143.86

Log likelihood = -697.05458   Prob > chi2   =   0.0000

```
--------------------------------------------------------------------------
         lnpcLOS |    Coef.  Std. Err.    z    P>|z|    [95% Conf.
Interval]
-----------------------+--------------------------------------------------
           age_t |  .0624101  .0003725  167.54  0.000    .06168   .0631402
           agesq |  .0004103  .0000239   17.16  0.000   .0003634   .0004572
       Mdivorced | -.0316301  .0122045   -2.59  0.010   -.0555505  -.0077097
  Mloneparentfams |  .0135823  .0051076    2.66  0.008   .0035716   .0235931
       Mregrepair | -.0101219  .0034573   -2.93  0.003   -.0168981  -.0033458
         Mcan_cit | -.0633706  .0185128   -3.42  0.001   -.099655  -.0270862
  Mfirst_generation | -.0321852  .0056304   -5.72  0.000   -.0432206  -.0211499
          Mab_iden |  .0025388  .0019618    1.29  0.196   -.0013063   .0063838
  Mem_participation |  .0278004  .0068757    4.04  0.000   .0143243   .0412764
          Mem_rate | -.0239102  .0057602   -4.15  0.000    -.0352  -.0126204
           Med_HS |  .0200474  .0044855    4.47  0.000   .0112559   .0288388
     Mvis_minority |  .0081282  .0051985    1.56  0.118   -.0020607   .0183172
      Mincome_fam |  3.87e-06  2.57e-06    1.51  0.131   -1.16e-06   8.90e-06
                  |
  c.Mdivorced#c.age_t | -.0010868  .0003531   -3.08  0.002   -.0017788  -.0003947
                  |
c.Mloneparentfams#c.age_t | -.0005157  .0001467   -3.52  0.000   -.0008032  -
.0002282
                  |
  c.Mregrepair#c.age_t | -.0000814  .0000962   -0.85  0.397    -.00027   .0001072
                  |
    c.Mcan_cit#c.age_t | -.0004875  .0004078   -1.20  0.232   -.0012869   .0003118
                  |
    c.Mab_iden#c.age_t | -.0000686  .0000539   -1.27  0.203   -.0001741    .000037
                  |
c.Mem_participation#c.age_t |  -.000212  .0001939   -1.09  0.274    -.000592
.0001681
                  |
    c.Mem_rate#c.age_t |  .0000971  .0001673    0.58  0.562   -.0002308    .000425
                  |
     c.Med_HS#c.age_t | -.0001987  .0001243   -1.60  0.110   -.0004424    .000045
                  |
  c.Mvis_minority#c.age_t |  -.000032  .0001233   -0.26  0.795   -.0002736   .0002096
```

```
                      |
     c.Mincome_fam#c.age_t |  1.74e-07  7.34e-08   2.38  0.018   3.05e-08   3.18e-07
                      |
              _cons |  -.0453973  .0124715   -3.64  0.000    -.069841  -.0209537
-----------------------------------------------------------------------------------------
```

```
--------------------------------------------------------------------------------
  Random-effects Parameters |   Estimate   Std. Err.    [95% Conf.
Interval]
-----------------------------+--------------------------------------------------
-
censusdivi~n: Unstructured   |
              sd(age_t) |  .0041703  .0003071    .0036099   .0048178
              sd(agesq) |  .0002527  .0000211    .0002145   .0002977
              sd(_cons) |  .1528443  .0092171    .1358059   .1720203
       corr(age_t,agesq) |  -.3764085  .1193857    -.5839828  -.1226313
       corr(age_t,_cons) |  -.2237343  .0906053    -.3923077  -.0406205
       corr(agesq,_cons) |  .0477786  .1213246    -.1882487   .2785883
-----------------------------+--------------------------------------------------
-
            sd(Residual) |  .2485517  .0019435    .2447716   .2523901
--------------------------------------------------------------------------------
LR test vs. linear regression:      chi2(6) =  2732.33   Prob > chi2 = 0.0000
```

Note: LR test is conservative and provided only for
reference.

Table A-10: Model 3 Regression for Males

Performing gradient-based optimization:

Iteration 0:   log likelihood = -1303.9416
Iteration 1:   log likelihood = -1303.9416

Computing standard errors:

```
Mixed-effects ML regression              Number of obs    =    8700
Group variable: censusdivision            Number of groups  =     174
```

```
                        Obs per group: min =        50
                                     avg =
50.0
                                     max =
50


                        Wald chi2(23)     =  34698.84
Log likelihood = -1303.9416              Prob > chi2     =    0.0000


-----------------------------------------------------------------------------
         lnpcLOS |    Coef.  Std. Err.    z   P>|z|    [95% Conf.
Interval]
-------------------------+---------------------------------------------------
           age_t |  .0692304  .0003847  179.97  0.000    .0684764   .0699843
           agesq | -.0000438  .0000279   -1.57  0.117   -.0000985   .0000109
        Mdivorced | -.0319985  .0097165   -3.29  0.001   -.0510425  -.0129545
   Mloneparentfams |  .0129444  .0040664    3.18  0.001    .0049744   .0209144
        Mregrepair | -.0072143  .0027526   -2.62  0.009   -.0126093  -.0018193
          Mcan_cit |  -.058817  .0147424   -3.99  0.000   -.0877117  -.0299224
   Mfirst_generation | -.0286307  .0044852   -6.38  0.000   -.0374216  -.0198398
          Mab_iden |   .001042   .001562    0.67  0.505   -.0020194   .0041033
   Mem_participation |  .0244874  .0054742    4.47  0.000    .0137583   .0352166
          Mem_rate |  -.022408  .0045859   -4.89  0.000   -.0313963  -.0134198
           Med_HS |  .0142555  .0035713    3.99  0.000    .0072559   .0212551
       Mvis_minority |   .007138  .0041396    1.72  0.085   -.0009754   .0152515
        Mincome_fam |  2.42e-06  2.04e-06    1.19  0.236   -1.58e-06   6.43e-06
                 |
   c.Mdivorced#c.age_t | -.0021225  .0003731   -5.69  0.000   -.0028539  -.0013912
                 |
 c.Mloneparentfams#c.age_t | -.0000557   .000155   -0.36  0.720   -.0003594
.0002481
                 |
   c.Mregrepair#c.age_t |   .000013  .0001017    0.13  0.898   -.0001862   .0002123
                 |
   c.Mcan_cit#c.age_t | -.0000312   .000431   -0.07  0.942   -.0008759   .0008135
                 |
   c.Mab_iden#c.age_t | -.0001624  .0000569   -2.85  0.004    -.000274  -.0000509
                 |
c.Mem_participation#c.age_t | -.0001456  .0002049   -0.71  0.477   -.0005472
.000256
                 |
     c.Mem_rate#c.age_t |  .0000536  .0001768    0.30  0.762   -.0002929   .0004001
```

|

c.Med_HS#c.age_t	-6.67e-06	.0001314	-0.05	0.960	-.0002642	.0002509

|

c.Mvis_minority#c.age_t	-.0000155	.0001303	-0.12	0.905	-.0002708	.0002398

|

c.Mincome_fam#c.age_t	1.79e-07	7.75e-08	2.31	0.021	2.69e-08	3.31e-07

|

_cons	.1297381	.0102773	12.62	0.000	.109595	.1498811

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]

-----------------------------+---

-

censusdivi~n: Unstructured |

sd(age_t)	.0042358	.0003158	.00366	.0049023
sd(agesq)	.0003068	.0000237	.0002637	.000357
sd(_cons)	.1206927	.0080509	.1059011	.1375502
corr(age_t,agesq)	.1091145	.1183645	-.1245843	.3313431
corr(age_t,_cons)	-.2986315	.0927487	-.4680728	-.1080103
corr(agesq,_cons)	-.1424413	.108168	-.3450488	.0728501

-----------------------------+---

-

sd(Residual)	.2677995	.002094	.2637267	.2719353

LR test vs. linear regression: chi2(6) = 1598.71   Prob > chi2 = 0.0000