Effect of Employment on Health-Related Quality of Life, Healthcare Expenditures, and Healthcare Utilization of Older Adults

Diana Kachan

University of Miami, dianakachan@gmail.com

Follow this and additional works at: https://scholarlyrepository.miami.edu/oa_dissertations

Recommended Citation
https://scholarlyrepository.miami.edu/oa_dissertations/1191

This Embargoed is brought to you for free and open access by the Electronic Theses and Dissertations at Scholarly Repository. It has been accepted for inclusion in Open Access Dissertations by an authorized administrator of Scholarly Repository. For more information, please contact repository.library@miami.edu.
EFFECT OF EMPLOYMENT ON HEALTH-RELATED QUALITY OF LIFE, HEALTHCARE EXPENDITURES, AND HEALTHCARE UTILIZATION OF OLDER ADULTS

By

Diana Kachan

A DISSERTATION

Submitted to the Faculty of the University of Miami in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Coral Gables, Florida

May 2014
EFFECT OF EMPLOYMENT ON HEALTH-RELATED QUALITY OF LIFE, HEALTHCARE EXPENDITURES, AND HEALTHCARE UTILIZATION OF OLDER ADULTS

Diana Kachan

Approved:

David J. Lee, Ph.D.
Professor of Epidemiology

Lora E. Fleming, M.D., Ph.D.
Professor of Epidemiology
University of Exeter Medical School

Guillermo Prado, Ph.D.
Professor of Biostatistics

Peter Muennig, M.D., M.P.H.
Professor of Health Policy and Management
Columbia University

Sharon L. Christ, Ph.D., M.S.
Professor of Statistics
Purdue University

M. Brian Blake, Ph.D.
Dean of the Graduate School
The proportion of adults aged 65 years and over in the U.S. population is rapidly increasing due to increasing longevity, as well as decreasing fertility rates. As a result, adults aged 65+ constitute the fastest growing group among workers, and this trend is projected to continue as the baby boomer generation enters retirement age. The positive effects of employment at older age on health has been suggested by previous research, however no comprehensive examination of health status of older workers has been conducted. In addition, the effect of employment at older age on such important health outcomes as health-related quality of life (HRQL), healthcare use, and healthcare utilization has not been well examined. Finally, the effect of occupation on health outcomes of older workers has been virtually ignored. This study fills this important gap in literature using nationally-representative data and complex statistical modeling techniques, such as structural equation modeling.

Using the data from the 1997-2011 National Health Interview Survey (NHIS) and the 2000-2009 Medical Expenditure Panel Survey (MEPS), we found that older workers are generally healthier than the non-workers, with those in more physically demanding occupations (farm, service, and blue collar workers) reporting the best health. Workers in
these occupations reported the fewest number of functional limitations and prior diagnoses of chronic illness, and had the highest mental health and overall HRQL scores. They were however just as likely as the white collar workers to rate their health as good or better, and had the same physical HRQL scores. Only service workers had lower healthcare expenditures than other working older adults, which could be explained by poorer healthcare access in this occupation. Unemployment at older age was however among the strongest predictors of increased healthcare costs, poorest HRQL, and poor health outcomes on all other measures.

In agreement with previous literature, we find a strong association between later life employment and improved health outcomes. Our results suggest a certain degree of healthy worker effect (e.g. in lower number of chronic health conditions in workers), as well as beneficial effect of employment on health at older age (e.g. improved mental HRQL scores among workers in physical occupations in the absence of differences between workers on physical HRQL scores). We discuss the implications of the results and provide suggestions for potential policy measures aimed at improvement of older adult work opportunities.
DEDICATION

I dedicate this dissertation to my parents, Vladimir Kachan and Aleksandra Gajewskaja, who always allowed me to do what I thought was right and never tried to change my mind, even when they did not at all approve of what I was doing. Even being on a different continent and not being able to see them for the first 6 years, I never stopped feeling their love. For what I can only imagine I have put them through in those years, this is for them.

And, of course, to the rest of my family: my brother Dmitry, who has been an endless source of inspiration; my brother Egor, whom I myself hope to someday inspire; my sister-in-law Alena, who has been inspiring both me and Dmitry; my nephew and niece, Timofei and Angela, who have been a tremendous source of joy; and, of course, Kisa.
ACKNOWLEDGEMENTS

I would like to thank my entire dissertation committee for their guidance, patience, and support. I could not have wished for a better group of people to guide me, and every single one of you has been instrumental to my success. Thank you for your expertise, mentorship, kindness, generosity, and for being always so accessible despite your incredibly busy schedules. I especially would like to thank Drs. David Lee and Lora Fleming for their incredible mentorship above and beyond any expectations I could have had. You have set a high standard, and I really hope I can emulate you someday.

I would also like to thank the NIOSH research group, who became my research family for a while, and specifically Kris Arheart, Manny Ocasio, Tainya Clarke, Alberto Caban-Martinez, and Noella Dietz. I have had an amazing time working (and not working) with you all. Thank you to my good friends Stacey Tannenbaum, Redinda Sherman, Dorothy Parker, Jacob Kabasso, and Isidoro (Izzy) Armenteros, your presence has made the graduate school experience a lot more cheerful. And lastly, I would like to thank Ms. Heather Rose for the incredible job she is doing at the Department of Public Health Sciences.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>LIST OF FIGURES</th>
<th>vi</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>vii</td>
</tr>
</tbody>
</table>

## Chapters

1. **Introduction and Background** .............................................. 1
2. **Study Methods** ................................................................. 19
3. **Specific Aim / Paper #1** ................................................... 54
4. **Specific Aim / Paper #2** ................................................... 68
5. **Specific Aim / Paper #3** ................................................... 85
6. **Discussion and Conclusions** ............................................... 101
7. **References** ................................................................. 138
8. **Appendix A: NHIS and MEPS Variables Used for Modeling** ........ 162
9. **Appendix B: Source Code** .................................................... 166
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figures</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.1: Population Pyramids by Gender and Year</td>
<td>2</td>
</tr>
<tr>
<td>Figure 1.2: Conceptual Model of the Study</td>
<td>12</td>
</tr>
<tr>
<td>Figure 2.1: Path Model Illustrating the Relationships Tested as a Part of Specific Aim 2</td>
<td>20</td>
</tr>
<tr>
<td>Figure 2.2: Path Model Illustrating the Relationships Tested as a Part of Specific Aim 3</td>
<td>21</td>
</tr>
<tr>
<td>Figure 2.3: MEPS-HC Overlapping Panel Design</td>
<td>25</td>
</tr>
<tr>
<td>Figure 2.4: Sample Sizes by Outcome - Specific Aim 1</td>
<td>28</td>
</tr>
<tr>
<td>Figure 2.5: Sample Sizes by Outcome - Specific Aims 2 and 3</td>
<td>29</td>
</tr>
<tr>
<td>Figure 2.6: Sources of Payment for Healthcare Expenses among Medicare Beneficiaries Aged 65+. 2007 Medical Expenditure Panel Survey</td>
<td>45</td>
</tr>
<tr>
<td>Figure 4.1. Full Path Model Predicting HRQL Outcomes</td>
<td>72</td>
</tr>
<tr>
<td>Figure 4.2: Standardized Path Estimates for Modeling HRQL Outcomes</td>
<td>77</td>
</tr>
<tr>
<td>Figure 5.1. Full Path Model Predicting Healthcare Expenditures and Utilization Outcomes</td>
<td>89</td>
</tr>
<tr>
<td>Figure 6.1: Conceptual Model of the Study, Updated</td>
<td>127</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Tables | Page
--- | ---
Table 2.1: Retirement Status by Employment/Occupation Group. Medical Expenditure Panel Survey Participants | 43
Table 3.1: Sample characteristics. National Health Interview Survey (NHIS) 1997-2011 participants aged 65+ | 59
Table 3.2: Multiple Logistic Regression Results for the Health Indicators. All National Health Interview Survey 1997-2011 Participants Aged 65+ | 60
Table 3.3: Multiple Logistic Regression Results for Multimorbidity by Gender. National Health Interview Survey 1997-2011 Participants Aged 65+ | 61
Table 4.2: Health Related Quality of Life: Regression Results and Select Path Model Results. Medical Expenditure Panel Survey 2000-2009 Participants Aged 65+ | 76
Table 4.3: Standardized Effects on Health-Related Quality of Life Through Mediated Pathways in the Final Model | 78
Table 5.1: Characteristics of the Sample. Medical Expenditure Panel Survey 2000-2009 Participants Aged 65+ | 91
Table 5.2: Zero-Inflated Negative Binomial and Negative Poisson Regression Results, Direct Effects | 92
Table 5.3: Mean Expenditures for Older Adult Subgroups. Medical Expenditure Panel Survey 2000-2009 | 96
Table 6.1: Distribution of Prior Diagnoses of Chronic Conditions in the MEPS Sample

Table 6.2: Predictors of Private Insurance Coverage as Compared to Public Coverage. Medical Expenditure Panel Survey

Table 6.3: Predictors of Having no Insurance Coverage as Compared to Public Coverage. Medical Expenditure Panel Survey

Table 6.4: Predictors of Employment and Occupation in Later Life as Compared to Being Employed in a White Collar Occupation. MEPS Data

Table 6.5: Multivariable Logistic Regression Results Modeling Multiple Functional Limitations. NHIS Data

Table 6.6: Multivariable Linear Regression Predicting the Number of Chronic Health Conditions. MEPS data
Chapter I: Introduction and Background

“Old age is the most unexpected of all the things that happen to a man.” Leon Trotsky

I. Overview

The US population is aging rapidly, with people aged 65 and over currently comprising 13.0% of the population, and their proportion expected to exceed 20% by year 2030 (Figure 1.1). In fact, this aging process is happening worldwide, as the life expectancy is increasing, and birth rates are dropping. In the US, the implications of our aging population include substantial increases in healthcare expenditures, and eventual unviability of the Social Security, Medicare, and the pension systems, as well as concerns about the elder quality of life and potential antagonism of the younger generations in the context of future shrinking resources and compromised economic wellbeing.

As a result of the global population aging process, and in order to meet the demands of financing the increased number of years lived in older age, the workforce is aging as well. Workforce aging is both a result of older adults constituting a larger proportion of the population, as well as them being better able and more willing to continue employment due to improved old age health as well as financial need. The number of workers 65+ has been increasing much more rapidly than any other worker age group, with the number of older workers projected to increase at five times the rate of the overall workforce. In addition, the older worker population is becoming more diverse, with a higher proportion of ethnic minorities and women than in the previous years.
Figure 1.1: Population Pyramids by Gender and Year.

Source: United States Census Bureau.¹⁰

If we are to generate proactive policies designed to address the needs of the working elderly, it is critically important to understand the distributional effects of the vast demographic changes on such socioeconomic health outcomes as health-related quality of life and resource utilization by the older workers. Specifically, occupational class, social status, race/ethnicity, and education are all likely to be important determinants of health among the current and future working elderly. It is important to understand how these changes impact health expenditures (one of the drivers of workforce participation among the elderly¹¹) and healthcare utilization, as such
knowledge could provide valuable insights into the burden of illness and healthcare resource needs of older workers, and provide guidance for the optimal allocation of limited resources in occupational safety and health.

II. Older Worker Characteristics

Before proceeding with the study description, let us define what exactly is meant by “older worker.” This term has been used in the literature both to describe those working past traditional retirement age (age 65 and older), as well as those who comprise the older group among the current working population depending, on the study. As a result, in the literature, the threshold age for “older worker” has ranged from 45 to 70. This study focused on the health effects of employment of the expanding older population, rather than on health issues in the workforce related to population aging. Health outcomes are highly dependent on health insurance coverage, and in that respect near-universal eligibility of adults aged 65 for Medicare makes this group very different from their younger “older” peers. In addition, the effects of work on health and wellbeing may change in the context of eligibility for retirement. Therefore, the population of interest in this study was limited to older adults for whom employment is to an extent a matter of choice and deliberate effort, and who with few exceptions have access to basic health insurance coverage, i.e. those aged 65+.

Employment past retirement age is different from that of younger individuals in that it is often not career employment. Older individuals might leave the job they held for a large portion of their life and that constituted career employment for a variety of reasons – being pressured to leave, seeking a position less demanding, whether physically
or mentally, seeking a more flexible schedule, or becoming eligible for retirement
benefits and therefore retiring from their previous job, but then returning to work in a
different capacity. Almost two-thirds of older workers who change jobs at older age also
change occupations, usually transitioning to lower-paying jobs that offer greater
flexibility. Older workers who change occupations tend to report greater job
satisfaction, despite taking a pay cut at a job with a lower social standing than their
previous career employment. Therefore the effect of employment on the health of older
individuals might be different from that of younger workers, and that specific association
between health and employment in later life is of interest in this study.

The benefits and drawbacks of older adult employment have been widely studied,
and yet the issue remains complex. The current older population is different from the
previous generation in that it is healthier and better educated, and therefore it is
better able to stay productively in the workforce past retirement age. However, the nature
of work has also changed over the lifetime of the current older worker from
predominately manufacturing-focused to predominately service-sector focused, with the
high-growth occupations either requiring special qualifications or offering low pay. With three quarters of older workers reporting making an income as a goal for seeking
employment, the complexity of the job tasks and lack of confidence in the skills required
can pose quite a problem and force the older workers to settle for lower paying jobs.
For some of older individuals, continued employment might be associated with health
benefits due to increased social interaction, mental and physical activity, increased
income, and access to better health insurance coverage; however, for others the
need to stay employed for the sake of making an income might be associated with health
detriments due to job stress, excessive physical exertion, harmful exposures, and risk of injury.  

From the employer’s perspective, declining physical functioning in older age could result in potential losses in productivity, need for greater schedule flexibility, and employee health insurance expenditures. Older worker’s skills might be out of date, and age-related changes in functioning might require special accommodations, such as better lighting; for example, older workers may require more lighting due to visual impairment. On the other hand, older workers are on average no less productive than younger workers, and if anything they are more careful and emotionally stable, demonstrate lower rates of absenteeism, and have developed more efficient strategies for performing work-related tasks.

Finally, from the policy perspective, an aging workforce will lead to higher overall rates of morbidity and disability among workers, however work might improve the health of the elderly population and provide easy access to the working elderly for health improvement measures. Therefore, if we are to encourage older individual to continue employment, it is important to identify groups that would be able to work safely and productively with the greatest benefit for their health.

III. Healthy Worker Effect

The relationship between health and employment is complicated by a selection bias termed a “healthy worker effect.” Individuals in poor health might find working difficult or uncomfortable, and as a result exit the workforce earlier. Their productivity might be reduced resulting in decreased wages, further discouraging employment. Employers
might also be less likely to hire workers in poorer health, especially if the job comes with a benefits package such that workers in poorer health cost employers more.\textsuperscript{24,42} In addition, sicker older adults who qualify for disability benefits on the condition that they do not work might avoid employment to preserve those benefits. As a result, poor health results in decreases not only in the amount of work hours and wages earned, but also in reduced likelihood of employment. For some individuals in poor health especially at the bottom of the socio-economic scale, the effect might be the opposite: increased healthcare costs associated with poor health might force them into employment.\textsuperscript{42} The healthy worker effect is strong among the older workers, possibly stronger than among the younger workers, and might even account for some of the better work-related health outcomes among older workers as compared to the younger workers (e.g. number of sick days).\textsuperscript{43,44}

\textbf{IV. Health Status of Older Adults}

Consistent with the healthy worker effect, as well as with the benefits associated with employment, older workers tend to be healthier, both mentally and physically, as well as have better survival rates than their non-working peers.\textsuperscript{12,13} However the better health of older workers is not all due to selection bias, as retirement was independently associated with worsening health, and employment with health benefits. In a longitudinal study based on the Health and Retirement Study (HRS), retirement was associated with decreased mobility, worsening mental health, and increased number of health conditions at six year follow-up.\textsuperscript{18} However, those who stayed active and had better social support did not fare as badly. Retirement before traditionally expected age was associated with
the worst health outcomes. A review of the literature on older adults’ engagement in either paid or volunteer work found such engagement to result in better wellbeing and mental health outcomes, lower mortality, and better ADL functioning. Another prospective HRS study found a dose-response relationship between the amount of work and health, and demonstrated this relationship to be bi-directional: as the individuals reduced the amount of time they worked, their health got worse, which led to a further decrease in the amount of work. A longitudinal study based on Asset and Health Dynamics among the Oldest Old (AHEAD) data found, however, that work beyond 100 hours per year was not associated with better outcomes compared to 100 annual hours.

In addition to potential effect of retirement on individual’s income and access to healthcare, retirement’s negative effects on health seem to operate through lifestyle changes such as decreased physical activity and social involvement. Such effects can be mitigated if the individual continues to work part-time or does volunteer work and engages in physical activity. However, retirement might also have a positive effect on health, at least in some older adult groups. For example, highly educated workers reported an improvement in health following retirement. Older British civil servants had decreased mental health functioning scores, while the scores of their retired peers, albeit only those with higher SES, improved post retirement. Elderly workers were also more likely to report unhealthy behaviors than retirees, and retirees are more than twice as likely to quit smoking as those who continue working.

A study based on longitudinal data from the German Socioeconomic Panel identified three trajectories of change in subjective well-being and life satisfaction following retirement, which can serve as a model for changes in health and well-being.
following retirement.52 Retirees experienced slight or dramatic increases or even dramatic drops in life satisfaction, and among the main factors predicting the direction of the effect were age at retirement, marital status, socio-economic status (SES), and physical health. The authors concluded that retirement improved wellbeing for those who were in poorer physical health by removing extraneous demands, did not make much of a difference for healthy married individuals of higher SES who had resources to fall back on in replacing their job role, and negatively affected wellbeing in those who did not have these resources but also did not benefit from no longer having a job. Among other factors, motivation for continued employment and the voluntariness of retirement play a major role in the actual direction of effect of retirement on health outcomes.14,18,53,54

Elderly workers are an increasingly heterogeneous population,55 and the needs for workplace accommodations may vary across subpopulations. Significant variation in health of older workers exists across occupational groups, such as by industry sectors or occupations characterized by different prestige levels.12,13 Elderly minority and female workers are at increased risk for occupational injury and death relative to other elderly workers.30,56 There are also differences in the risk of certain conditions (e.g. arthritis, back problems, chronic lung disease, functional limitations, work disability, and work-related injuries) across the occupations of older workers.57,58 However, there has been little research on the health of different race-ethnic, gender, and socio-economic subpopulations of the working elderly, despite clear evidence of health disparities among select subgroups.

Health is a complex concept to characterize - something that everyone has a pretty clear idea about, and yet it is quite difficult to measure.42,59 Currie and Madrian (1999)
identified eight health measures relevant to “work capacity” and associated with labor force participation: 1) self-rated health status; 2) limitations in the ability to work due to health; 3) other functional limitations (e.g. ADLs); 4) presence of health conditions, both chronic and acute; (5) health care utilization; (6) mental health and substance use assessment; (7) nutritional status; and (8) expected mortality, with the first 5 of these measures being most relevant to developed countries. This study examined the first 4 of these measures in chapter 3, while chapter 5 also examined health care utilization.

V. Health-Related Quality of Life in Older Adults

Health-related quality of life (HRQL) is an important measure of health that is predictive of mortality, healthcare costs, healthcare service use, and other health outcomes, as well as earlier retirement, and its improvement is one of the overarching goals of the Healthy People 2020. HRQL measures evaluate both the physical aspect of ill health as well as its effect on overall emotional or social functioning, as it is perceived by the individual. Generally, people reporting poorer HRQL are more likely to be female, older, racial/ethnic minorities (except Asian/Pacific Islanders), individuals with lower education and income, and those with a greater number of chronic disease or disability. Between 26.0% and 33.2% of older adults in the US currently report being in fair or poor health compared to 7.3-14.4% among adults aged less than 55. Yet, a large portion of the adults aged 65+ also describe their current age as the “best years of my life.”

Meaningful social participation, such as participation in the workforce, may be associated with improved HRQL, while social isolation is associated with lower HRQL.
Physical activity is also associated with better quality of life in the elderly;\textsuperscript{71} retirement may negatively affect individual’s health by reducing the amount of physical activity.\textsuperscript{18,72} HRQL varies across income levels among adult workers\textsuperscript{25}, however it is unknown whether such variation exists among older workers. To date, a systematic comparison of quality of life measures in worker subgroups, and in particular in workers 65+, has not yet been undertaken in the US. The inclusion of two well-known quality of life measures, Short Form-12 (SF-12) and EuroQol-5D (EQ-5D), in the Medical Expenditure Panel Survey (MEPS), one of the databases used in this study, provided an opportunity to comprehensively address this gap in the literature.

VI. Healthcare Expenditures and Utilization in Older Adults

Adults aged 65+ are the age group with the highest healthcare costs.\textsuperscript{73,74} Up until the beginning of 1990s, employer retirement plans often included health insurance coverage, however due to rising healthcare costs those have been getting increasingly rare.\textsuperscript{75,76} Since the employer contributions to retiree health insurance has decreased substantially in the last 2 decades,\textsuperscript{77} so did the research of what happens to the healthcare costs and utilization as the person enters retirement in the United States. The healthcare expenditures and health services use of older adults have been widely studied in Europe and Canada. The findings indicate that socio-economic and gender differences exist in a number of healthcare use indicators, such as the number of hospital visits, medications prescribed, specialist visits, and overall healthcare costs.\textsuperscript{78-82} Women used more health care services overall.\textsuperscript{81,83} More specifically, women were less likely to be admitted to a hospital, see a specialist, or go to the emergency room; however they were more likely to
see a family physician than men.\textsuperscript{81,82,84} Greater numbers of physician and emergency room visits were also associated with greater age within the elderly group, possibly due to an increased number of medical conditions.\textsuperscript{79,85-87} Chronic conditions also accounted for most of the differences in hospitalization rates between the retired and working Canadian older adults.\textsuperscript{88} In the US, retired older adults were more likely to utilize emergency room or urgent care services, though this increase in utilization was need-based and reduced by factors that improved primary care access.\textsuperscript{89,90} Retired older adults were also more likely to see a physician if they were previously self-employed, as the self-employed individuals would have been less willing to take time off work prior to retirement.\textsuperscript{91} Retired older adults were less likely to seek dental care, and this effect was mediated by decreased family income, lack of dental coverage, however also by increased free time available.\textsuperscript{92} However, overall there is a lack of studies comparing healthcare use or expenditures between the retirees and the employed older adults in the US over the last 20 years, and there are none comparing these outcomes within older worker subgroups. Through increased income, access to better insurance coverage, positive effects on health through social interaction, and through putting higher demands on older adults’ time, employment could potentially have an impact on older workers’ healthcare costs and service use.

\textbf{VII. The Overall Conceptual Model of the Study}

The conceptual model of this study (Figure 1.2) is based on the biopsychosocial model of health, which postulates that the individual’s health is affected by a combination of biological (e.g. gender, co-morbidity, genetic vulnerability, etc.),
psychological/behavioral (e.g. stress coping skills, unhealthy behaviors), and social (e.g. education, health insurance, employment, income, etc.) factors. This conceptual model builds on the biopsychosocial model to incorporate the widely used existing models of HRQL and healthcare utilization.

Figure 1.2: Conceptual Model of the Study

Among the most commonly used conceptual frameworks of HRQL is the causal pathway model developed by Wilson and Cleary, and further revised by Ferrans et al. This model proposes the following sequential steps to the formation of one’s HRQL status: 1) biological function; 2) symptoms of ill health; 3) functional status; 4)
perception of one’s health; and 5) health-related quality of life. Each of these individual steps is in turn affected by characteristics of the individual (demographic, biological, psychological, and developmental), as well as by the physical and social characteristics of the individual’s environment (e.g. neighborhood and workplace safety, cultural influences, etc.).

That is, this model presents HRQL as an outcome of a sequence of health status parameters, all of them in turn affected by the biopsychosocial model’s predictors of health. While information on the environmental factors was not available through the data sources used in this study, information was available on a variety of individual factors. In addition, the health status factors included in the causal pathway of HRQL were measured as a part of the National Health Interview Survey (NHIS) in the form of the following: prior lifetime diagnosis of chronic health conditions; functional limitation status; self-rated health; and information on activity limitations; this allowed the calculation of the Health and Activity Limitation index (HALex, a measure of HRQL).

The widely used Andersen and Newman model of health service utilization proposes a sequence of the following factors in predicting healthcare utilization: 1) predisposing factors, which characterize the propensity of certain groups to use more healthcare and include such factors as demographics, beliefs about the healthcare system, and social factors such as education, occupation, race; 2) enabling factors, which characterize the resources available to the individual through family or community and include income and health insurance coverage; and 3) illness level, both perceived and as evaluated by a healthcare practitioner, affecting the perception of the need for healthcare.

In combination with the health system and societal factors, such as
technology available and health system organization, all these affect the type, purpose, and frequency of healthcare use.\textsuperscript{98} Again the data sources used in this study allowed only the inclusion of individual level predisposing and enabling factors, as well as illness level, represented by the number of chronic health conditions.\textsuperscript{85}

The conceptual model in Figure 1.2 is a hybrid of the three models described above to the extent that the data were available through the NHIS and the MEPS, the two study databases. Individual factors predicting health outcomes were subdivided into the relatively constant (at older age) socio-demographic variables and health behaviors, which would not be predicted by other variables in the model (e.g. education, age, gender, race/ethnicity, smoking, and alcohol consumption); factors more prone to change, which were placed as mediators (e.g. health insurance status, income, employment/occupation).

Three sub-models, one for each of the 3 chapters, were developed based on the general model in Figure 1.2. The effects of the individual characteristics on the HRQL causal pathway components (e.g. chronic illness, functional limitations, self-rated health, HRQL/HALex; top part of Figure 1.2) were modeled using NHIS data, and described in chapter #3. The use of the NHIS data allowed for a greater sample size and inclusion of more detailed health behavior information. For chapter #4, detailed information on the HRQL measures was obtained through the MEPS, and modeled as a function of individual characteristics and chronic health conditions while controlling for the effects of mediation pathways. The causal pathway from chronic illness to HRQL was not included in the chapter #4 sub-model; only the direct pathway from chronic illness to HRQL was retained in order to increase the statistical power available for structural
equation modeling (SEM). The effects of individual characteristics and chronic health conditions on healthcare expenditure and utilization outcomes were examined using MEPS data in chapter #5. The use of the MEPS data allowed access to more detailed HRQL, expenditure, and health service use data; however it resulted in decreased sample size and limited health behavior information.

VIII. Dissertation Objectives

Using the pooled data from the 1997-2011 National Health Interview Survey (NHIS) and the 2000-2009 Medical Expenditure Panel Survey (MEPS), this study explored the aforementioned associations through the following specific aims and hypotheses:

**Specific Aim 1.** Provide comprehensive descriptive statistics on US workers and non-workers aged 65+ to evaluate measures of health disparities, co-morbidities, and health-related quality of life (HRQL) using nationally representative NHIS data pooled from 1997-2011 (n= 86,454). Occupation was categorized based on the National Center for Health Statistics (NCHS) classification.

**Specific Aim 2.** Using SEM and the 12-Item Short-Form Health Survey (SF-12) data available in 2000-2009 MEPS (n= 34,643), explore the quality of life of older worker and non-worker subpopulations. SEM techniques were used for the prediction of quality of life in different older worker sub-populations after adjustment for health behaviors, co-
morbidity, and sociodemographic characteristics. The conceptual model underlying the SEM analyses is presented in Figure 1.2.

$H_{2.1}$: While workers in white collar occupations who continue working past retirement age will demonstrate higher quality of life relative to retired peers, work past retirement age for workers in other occupations will result in lower quality of life relative to their retired peers after adjustment for other factors (MEPS).

$H_{2.2}$: While older workers in race/ethnic minority subgroups will experience lower quality of life than non-Hispanic white older workers, this effect will be eliminated by introducing a mediation effect via occupation (MEPS).

**Specific Aim 3.** Using SEM, compare health care utilization and expenditures in working and non-working US adults 65+ (n=34,643). These outcomes controlled for health insurance coverage (public/private/none) and the number of co-morbidities (Figure 1.2).

$H_{3.1}$: Total healthcare expenditures will be significantly lower in older workers versus older non-workers, even after adjustment for sociodemographic factors and co-morbidities (MEPS);

$H_{3.2}$: Total healthcare expenditures will be significantly higher in blue collar than in white collar older workers, and this relationship will be only partially mediated by the type of health insurance coverage (MEPS).
This study is relevant to the research priorities of several organizations, including the National Institute on Aging, which provided support for the authors’ research and training program, and the National Institute of Occupational Safety and Health. These include: understanding the patterns of disease and disability in older adults, supporting the development of interventions to improve their health and quality of life, identifying health disparities and interventions to decrease these disparities, and understanding the implications of an aging society. The results from this study will be helpful to policy makers, employers, and other stakeholders interested in making health-related and employment-related decisions, and implementing workplace changes and health promotion programs targeting the senior community in order to improve the health, well-being, and quality of life of people 65+.

IX. Summary

Older US workers represent a large and growing proportion of the population that remains largely understudied. This group also presents a great potential for targeted interventions that could improve these workers’ wellbeing, both from the economic and from the health perspective, with implications for their families, older non-workers, and society in general. Multiple studies have examined the effects of retirement on health status, as well as healthcare utilization and expenditures, of the elderly population. While previous research demonstrated that continued employment affects the health of elderly individuals in ways that vary across various socio-demographic groups, there has been little exploration of the occupational factors affecting the health status and health resource utilization of older workers. Publically available data allow exploration into these relationships, as well as characterization
of the health status of this important worker group. The underexplored associations of employment and occupation with health, quality of life, and healthcare expenditures of the elderly are essential to understanding the impact of the growing older worker population on the current and future healthcare and public health systems.

Using large nationally representative worker population data, this study provides a unique examination of health and quality of life data by demographic, socio-economic, and occupational groups. Results help to highlight the importance of needed changes in retirement, health insurance, and employment policies, as well as promote a nuanced awareness about the risks and benefits of working into older age. This research will assist with the further development of workplace strategies and health-promotion programs to meet the unique needs of the elderly workers, their families and society.
Chapter II: Study Methods

I. Overview

The purpose of this chapter is to provide additional details on the different methodologies used in this dissertation. This more detailed information is meant to complement the methodology sections of the specific paper chapters without too much redundancy, while still providing the essential information in one place. This chapter: describes the overall study design, data sources, and sample selection process; provides details on study variables and data preparation; and describes the statistical analyses methods used for modeling. Appendices A and B supplement this chapter with a list of variables used and the source codes used for analyses.

II. Study Design

This research is based on secondary data analysis. The data publically available through the National Health Interview Survey (NHIS) and the Medical Expenditure Panel Survey (MEPS) were collected, de-identified, and coded by the National Center for Health Statistics (NCHS) and the Agency for Healthcare Research and Quality (AHRQ), respectively, as stipulated by the National Health Survey Act of 1956. The characteristics of each survey are described in the Data Sources section of this chapter.

The overall Conceptual Models of this dissertation was presented in Figure 1.2 and was described in chapter 1. On the top of the diagram are the outcomes of interest in specific aim 1, and on the right side are the outcomes of interest in specific aims 2 and 3. On the left and bottom of the diagram are the health disparities and other risk factors that appear to
influence these outcomes. The directionality of the relationships is not meant to necessarily imply the direction of causation, but rather statistical mediation effects. Sub-models for chapters 4 and 5 (Figures 2.1 and 2.2) were created on the basis of the overall conceptual model for each of the outcomes in the study: health-related quality of life (e.g. SF-12 physical and mental health component summary scores, EuroQol-5D (EQ-5D), healthcare expenditures, emergency room visits, and nights of hospital stay. These sub-models were used to create structural equation path models used for the statistical analyses.

Figure 2.1: Path Model Illustrating the Relationships Tested as a Part of Specific Aim 2
III. Data Sources

i. National Health Interview Survey (NHIS)

First administered in 1957, the NHIS is an annual multistage probability survey that is designed to be representative of the United States civilian non-institutionalized population living in addressed dwellings. The survey is continuous, with participants being interviewed throughout the year. Data are collected during face-to-face interviews using Computer Assisted Personal Interviewing (CAPI), with interviewers trained by the US Census Bureau according to the NCHS guidelines. Information is collected about socio-demographic and health-related characteristics (e.g. chronic illness, disability, health insurance coverage, smoking history, etc.) of the participants. In addition, supplemental information on cancer diagnosis history,
alternative and complementary medicine use, etc. is collected in some of the survey years.

The NHIS underwent a major redesign in 1997. Since then, one randomly selected adult aged 18 years and older reports information for the entire household. The redesign greatly improves the value and reliability of the health information collected compared to pre-1997 design, however it precludes pooling the data collected after 1997 with that collected prior to redesign.

In order to provide precise estimates for minority subpopulations of the US residents, the NHIS oversamples Black, Hispanic, and starting year 2006, Asian populations. Sampling weights are supplied and were used in order to obtain unbiased estimates for the population. The annual NHIS sample adult core sample size has varied from 21,781 to 36,116 over the years of this study. The average annual response rates for the 1997-2011 NHIS sample adult core were 70.3% (range: 61%-80%).

The current design of the NHIS includes the Core components and Supplements. The major Core components are Household (contains information on household composition), Family (socio-demographic characteristics of the family), Person, Sample Adult, and Sample Child (all 3 contain information on individual level health indicators: demographic characteristics, health status, health behaviors, chronic conditions, etc.). From each family, one adult and one child are randomly selected, for whom the information is collected for the Sample Adult and the Sample Child components. Starting in 2006, adults aged 65+ who are black, Hispanic, or Asian have a higher probability of being selected for the sample adult component.
The Family Core component serves as a sampling frame for other surveys integrated with the NHIS, such as the MEPS. In addition to Core components, supplements are issued periodically to provide information on particular areas such as cancer screening, complementary and alternative medicine use, children’s mental health, etc.

For Specific Aim 1, information was gathered from the core components of the NHIS year 1997-2011 with the focus on adults aged 65+. The variables of interest were drawn from the following component files: Persons (self-rated health, Health and Activities Limitations index components, education level, race/ethnicity, sex, age); and Sample Adult (lifetime diagnosis of chronic conditions, functional limitations, smoking and alcohol consumption history, employment and occupation). Sample Adult file sampling weights were used to obtain nationally representative estimates.

ii.  Medical Expenditure Panel Survey (MEPS)

Starting in 1996, the sampling frame for the MEPS was selected from the households which participated in the NHIS in the previous survey year. Like the NHIS, the MEPS is designed to be representative of the US civilian non-institutionalized population, and it collects a variety of information on its participants’ health service use, costs, sources of payment for healthcare, health status, health insurance coverage, as well as demographic, social, and economic characteristics. In addition to the groups oversampled as a part of the NHIS, the MEPS oversamples additional policy relevant sub-groups such as low income households, and it has been oversampling Asians since its inception. Each year’s sample size and subgroups
oversampled depend on the eligible sample available from the NHIS, and current DHHS objectives, as well as the budget available at the time of sampling. The sample size for the 2000–2009 MEPS household component ranged from approximately 23,800–37,400 individuals (full year data available) with response rates of approximately 57.2-66.3%.

Five in-person interviews, one during each of the 5 rounds, are completed over a two-year period for each MEPS panel using the CAPI (see Figure 2.3 for an illustration of the MEPS overlapping panel design). A new panel of households is randomly drawn from the previous year’s NHIS each year.105 The MEPS data are initially collected through individuals. After obtaining signed releases from MEPS respondents, expenditure data are also collected directly from hospitals, physicians, pharmacists and home healthcare providers; and insurance data are collected from the employers.106

The survey currently consists of three components: the Household Component (MEPS-HC); Medical Provider Component; and Insurance Component. Medical Provider Component is currently only used to edit and impute the MEPS-HC data and is not available as stand-alone data. The MEPS Insurance Component contains information collected through employers about employer-based health insurance and is not available for public use. The MEPS-HC is the only MEPS component currently available for public use, and it is the component used for analysis in this research.

The MEPS-HC component data are collected during household interviews. A household can contain more than one family unit, each consisting of more than one individual; however, only one individual is selected to report the data for the entire
The core MEPS-HC interview, which is administered during each round, collects information on the individual’s demographics and health status, specific medical conditions, healthcare expenses charged and paid, healthcare utilization, insurance coverage, and current employment. In addition, once a year the following are assessed using CAPI: Access to care, Child preventive health, Satisfaction with health plans & providers, Income, and Preventive Care.

During round 5 CAPI is also utilized to assess household assets. In addition, 2 paper questionnaires are administered to all adults by mail once a year during rounds 2 and 4: 1) Diabetes Care Survey (administered to participants with diabetes) and Adult Self-Administered Questionnaire (SAQ). The SAQ evaluates the participant’s view of healthcare access and quality of care (via Consumer Assessment of Health
Plans questionnaire), their health attitudes, and their health-related quality of life via 2 measures, SF-12 and EQ-5D. In 2003, SF-12 measure was replaced with SF-12v2. EQ-5D was only administered through year 2005.

Health information for each individual is collected and stored from two perspectives: for every medical condition reported, and for every event of health service utilization reported. These are stored as event-level and condition-level data files in addition to person-level and job-level (e.g. wages, hours worked, etc. for each job reported for the participant) files. These are available through the MEPS website as either point-in-time files (data for the beginning of the year) or full-year files. The full year files include: 1) Full Year Consolidated Data File (person level); 2) Event File; 3) Medical Conditions File; 4) Jobs File; 5) Person Round Plan Public Use File. This research used the data from the Full-Year Consolidated File.

Because most participants of the MEPS-HC contribute to two consecutive years of data, MEPS-HC samples of different years are not completely independent. Nevertheless, the variance structure of data files prior to year 2002 was calculated independently of the other years. Therefore MEPS-HC files cannot be directly pooled with each other, however proper variance structure for pooling of data over multiple years is provided in MEPS file HC-036. Therefore, for this research, variance information from the HC-036 was utilized to obtain nationally representative estimates using the pooled MEPS data. SAQ weights were used for all analyses, and these were adjusted for 4 years of pooled data when modeling EQ-5D, and for 10 years of pooled data when modeling all the other outcomes.
IV. The study population

The sample selection procedure with the corresponding sample sizes is illustrated for each dataset in Figures 2.4 and 2.5. The NHIS and MEPS variables and codes used for analyses are presented in Appendix A.

For chapter 3 analyses, the NHIS data were pooled for years 1997-2011 (Figure 2.4). Participants were included in the analyses if they were aged 65 or older and had employment and occupation (for employed) information available (n=86,454). Observations were excluded if they were missing information on any of the predictor variables (n=2,910). No participants were missing race/ethnicity, sex, or age information; therefore observations excluded were missing information on education level, smoking history, and alcohol consumption. Next, for each of the outcomes modeled in paper 1, observations were excluded if they were missing data on that outcome. The final sample sizes by outcome were the following: multimorbidity, n=83,521; multiple functional limitations, n=83,338; fair/poor self-rated health, n=83,457; low HALex, n=83,457.

For Specific Aims 2 and 3, the MEPS data pooled over years 2000-2009 were used (Figure 2.5). Here the only inclusion criterion was aged over 65 years (n=34,643). Due to full-information maximum likelihood treatment of missing data in Mplus, no observations with missing data needed to be excluded for modeling most outcomes with the exception of EQ-5D. Because EQ-5D was only measured in years 2000-2005, data for years 2006-2009 were excluded from analyses for this outcome (resulting sample n=20,795).
Figure 2.4: Sample Sizes by Outcome - Specific Aim 1

Specific Aim 1
NHIS survey years 1997-2011
Age 65 and older
Employment/occupation information available
n=86,454

Excluded (n=2,910)
- Missing education (n=1,191)
- Missing smoking status (n=871)
- Missing alcohol consumption (n=1,867)

Outcome: multimorbidity
Excluded (n=23)
- Missing prior diagnosis information
n=83,521

Outcome: multiple functional limitations
Excluded (n=206)
- Missing functional limitation information
n=83,338

Outcome: fair/poor self-rated health
Excluded (n=87)
- Missing self-rated health information
n=83,457

Outcome: low HALex
Excluded (n=87)
- Missing HALex score information
n=83,457
Figure 2.5: Sample Sizes by Outcome - Specific Aims 2 and 3

- Specific Aim 2 & Specific Aim 3
  - MEPS survey years 2000-2009
  - Age 65 and older
  - n=34,643

- SF-12 Physical and Mental Components
- Total Healthcare Expenditures
- Number of ER visits
- Nights of Hospital Stay
  - Excluded (n=0)
    - Information available on less than 2 variables
  - n=34,643

- EQ-5D
  - Excluded (n=20,827)
    - Years 2004-2009 EQ-5D was not measured
  - n=13,816
V. Study Variables

The following section describes the creation of variables used in the analyses. The specific questions asked of the NHIS and MEPS participants to collect information have changed over the study years, and survey questions, where quoted, are meant to be a generalized form of the questions asked over the years. Unless otherwise noted, if the original NHIS/MEPS variable was coded as “Not Applicable”, “Not ascertained”, “Don’t Know”, or “Refused”, the generated variable was coded as missing. The NHIS and MEPS variables are presented in Appendix A; the SAS codes used for the analyses are presented in Appendix B. The variables created are described by the dataset, with outcome variables described first and followed by predictor variables.

i. NHIS variables

a. Multimorbidity

Multimorbidity was defined as the self-reported concurrent presence of two or more chronic conditions, which were previously diagnosed by a health care professional. Chronic disease diagnoses were assessed annually with a question: “Have you ever been told by a doctor or other health professional that you had…” (yes/no). This was asked about the following conditions during all years of this study: hypertension, stroke, emphysema, asthma, cancer, coronary heart disease, angina, myocardial infarction, and diabetes. For the multimorbidity variable in this research, a heart disease variable was created that reflected whether the individual was ever diagnosed with any of the following: coronary heart disease, angina, myocardial infarction. The multimorbidity variable was then calculated as the sum of the lifetime
diagnoses of the following conditions: hypertension, stroke, emphysema, asthma, cancer, heart disease, and diabetes.

b. Multiple functional limitations

Presence of functional limitations was assessed with the question, “By yourself, and without using any special equipment, how difficult is it for you to...” with regards to the following activities: walking a quarter of a mile, climbing 10 stairs, standing for 2 hours, sitting for 2 hours, stooping or kneeling, reaching over one’s head, grasping small objects, carrying heavy objects, pushing large objects, going out for things like shopping, participating in social activities, and relaxing at home. Participants responded with a rating of their ability to perform these on a five-point Likert scale from ‘not at all difficult’ to ‘can’t do at all’.

This research defined a functional limitation as any difficulty (i.e. any answer other than ‘not at all difficult’) reported by the participant. This broad definition was chosen in order to capture the greatest number of older adults afflicted by such limitations. The number of activities on which each participant reported limitations was summed. Multiple functional limitations were defined as more than one limitation, and the participants were dichotomized into those who reported one or no limitations, and those who reported two or more.

c. Self-rated health

Participants rated their perceived health in response to the question: Would you say [your] health in general is excellent, very good, good, fair, or poor?” These
were dichotomized as 1) fair/poor and 2) good or better, as is common NCHS practice.\textsuperscript{111}

d. Health and Activities Limitation Index (HALex)

The HALex, is a health–related quality of life measure developed specifically for use with the NHIS data.\textsuperscript{112} It was calculated as described by Livingston and Ko.\textsuperscript{113} HALex is a utility score, and as such it combines information about an individual’s perceived health and activity limitations into a single index score that ranges from 0 to 1. The utility index value represents how preferable a certain health state is for the individual, with a score of 1 representing perfect health, and a score of 0 representing a health state equivalent to death.

The HALex combines information on the individual’s perceived health (i.e. self-rated health) with information on activity limitations. The activity limitations information is different from the functional limitations described above in that it incorporates the mental and emotional health aspect. The activity limitations are reported in response to the questions, which ask if due to a physical, mental, or emotional problem the person: 1) needs help of other people with personal care needs (eating, bathing, dressing, or getting around inside the home), 2) needs help of other people with routine needs (everyday household chores, doing necessary business, shopping, getting around for other purposes), 3) is kept from working at a job or business, 4) limited in the kind or amount of work they can do, or 5) limited in any other way in any activities.\textsuperscript{114} The odds of low HALex index were modeled in paper
1, and low HALex score was defined as below 20th percentile for the population (HALex<0.48) in order to allow a sufficient sample size in the low score population.

e. Employment / occupation

The definition of employment included both paid and unpaid (e.g. working at a family business or farm) work, both full time as well as part time. Employment and was based on the question: “Which of the following [were you] doing last week?” Participants who chose the following answer options were classified as employed: ‘Working for pay at a job or business”, “With a job or business but not at work”, “Working, but not for pay, at a family-owned job or business”. Employment was coded as a dichotomous yes/no variable.

The participants who reported working were asked further questions about their jobs. Participants’ verbatim responses were recorded about what their job/occupation was during the previous week. US Census Bureau coding specialists later review this information and assign occupation codes consistent with Standard Occupation Classification (SOC). The SOC classification changed in 2002, and the NHIS coding of occupations changed accordingly in 2004 consistent with the 2000 New SOC coding system. Based on the SOC Occupation Subgroups and Major Occupation Groups, the NHIS further recodes the occupation information into 41 and 13 categories prior to 2004, and into 94 and 23 categories starting in 2004 and after respectively.

For this research, the occupational categories provided by the NHIS were further condensed into 4 groups according to Krieger et al, with one added category
for those not in the workforce at the time of the interview: 1) unemployed/retired; 2) white collar, 3) blue collar, 4) farm worker, and 5) service worker. There was no distinction made between those participants who had retired, and those who were unemployed for another reason.

\textit{f. Race/ethnicity}\n
The NHIS participants report their self-identified race and Hispanic origin information in response to several questions. Individuals are asked to select one or more racial groups that they consider themselves to belong to. Next they are asked to select one group that best represents their race. Hispanic origin was assessed with the question: “Do [you] consider [yourself] Hispanic / Latino?” Individuals who respond affirmatively are asked to identify a specific group that represented their Hispanic ancestry. The information obtained with these three questions was used in this dissertation to create the following race/ethnicity categories: non-Hispanic white [reference], non-Hispanic black, Hispanic, other.

\textit{g. Sex}\n
Sex was classified as 1) male; 2) female.

\textit{h. Age}\n
Age at the time of interview was treated as a continuous variable. The NHIS top-codes age at 85 years; that is all ages greater than 85 are recorded as 85.
i. *Education level*

Education level was based on the question: “What is the HIGHEST level of school [you] completed or the highest degree [you] received?” This was classified as: 1) less than high school (including those who completed 12 years of education but did not get a high school diploma) [reference group], 2) high school (including GED or equivalent), and 3) more than high school (>12 years of education).

j. *Smoking history*

Lifetime smoking as well as current smoking was assessed in the NHIS. If the person reported ever smoking 100 cigarettes, they were asked if they currently smoked. Smoking history was classified as: 1) never smoker (less than 100 cigarettes in a lifetime), 2) current smoker (100 cigarettes in a lifetime and currently smoke every day or some days), and 3) former smoker (currently smoke less frequently than some days, but over 100 cigarettes over lifetime).

k. *Alcohol consumption*

Participants were asked about the amount and frequency of current drinking if they reported having had at least 12 drinks in their lifetime. This information was classified as: 1) never [reference] drinker (less than 12 drinks in a lifetime), 2) former drinker (no alcohol in the last year), 3) current light (drank in the past year, but never had 5 or more drinks a day), and 4) current heavy (reported drinking 5 or more drinks in a day on at least one occasion within the last year).
ii. MEPS variables

a. Health-related Quality of Life

Starting in year 2000, the Self-Administered Questionnaire (SAQ) has been mailed to all adult (age 18+) participants of the MEPS during rounds 2 and 4. The SAQ contains two measures HRQL, the EuroQol 5-D (EQ-5D), which was a part of SAQ through year 2003, and the Short-Form 12 (SF-12), which was replaced in 2003 with a Short-Form 12 Version 2 (SF-12v2). Both instruments provide valid and reliable assessments of the quality of life, including that of older people. SF-12 scores from years 2000-2002 of the MEPS data were recoded into SF-12v2 scores according to the SF-12 manual.

**EuroQol (EQ-5D)**, like the HALex described above, is a utility index. As such, it varies from 0 to 1, with higher scores representing health states that are more preferable to the individual. The EQ-5D is a widely based generic HRQL measure that based on five questions, which ask the participants to rate of a 3-point Likert scale how much they experience the following on that day: 1) problems with mobility, 2) problems with self-care, 3) problems with usual activities, 4) pain or discomfort, and 5) anxiety/depression. With 3 possible choice options on 5 dimensions of health, EQ-5D results represent 243 distinct health states, each of which has a score associated with it. The scoring systems vary between different countries, however the MEPS uses a scoring algorithm specific to the US preferences. In addition, the participants are asked to rate their health on a scale from 0 to 100. The responses to the first five questions are used to calculate the utility index, which is available in the MEPS as the EQU42 variable.
Therefore, EQ-5D generates two scores, a utility index and a rating score. Of these two, the utility index was used in this dissertation. Of note, while all utility scores are calculated to fit into the range between 0 to 1, negative EQ-5D scores are possible, and signify health states to which death is preferable by the individual. Because EQ-5D was only assessed during years 2000-2003, only these years of the MEPS data were used modeling it.

**SF-12** is a health profile measure of HRQL. It consists of two calculated scores, physical component summary (PCS) and mental component summary (MCS) scores, each characterizing a different aspects of health. SF-12 is based on 12 questions assessing the following information: 1) general health on the day of the survey; 2) during a typical day, limitations in moderate activities; 3) during a typical day, limitations in climbing several flights of stairs; 4) during past 4 weeks, as result of physical health, accomplished less than would like; 5) during past 4 weeks, as result of physical health, limited in kind of work or other activities; 6) during past 4 weeks, as result of mental problems, accomplished less than would like; 7) during past 4 weeks, as result of mental problems, limited in kind of work or other activities; 8) during past 4 weeks, pain interfered with normal work outside the home and housework; 9) during the past 4 weeks, felt calm and peaceful; 10) during the past 4 weeks, had a lot of energy; 11) during the past 4 weeks, felt downhearted and depressed; and 12) during the past 4 weeks, physical health or emotional problems interfered with social activities.

The calculation of both PCS and MCS scores involves information from all 12 questions, however the questions are weighed differently in the two formulas. In
the MEPS, calculated PCS and MCS component summary scores are available as variables PCS42 and MCS42. The algorithm of SF-12 scoring is created so that the score ranges from 0 to 100 with a mean of 50 and a standard deviation of 10 in a general population, with higher scores representing better HRQL. However, negative scores are possible in rare cases.

b. Healthcare Utilization and Expenditures

Information about healthcare service utilization and expenditures associated with utilization events is collected during every round through the Household component of the MEPS. Expenditures are defined in the MEPS as the amount that was actually paid from all sources for care provided. Data are collected on the following types of events: office-based and hospital outpatient medical provider visits; hospital inpatient stays and the number of nights per stay; emergency room visits; dental visits; prescription medicine purchases; home health care; and other medical equipment and health care services utilization events. Office-based and outpatient visits include both visits to physicians and non-physician providers (e.g. chiropractors, nurse practitioners, optometrists, etc.). Over-the-counter medicine, complementary/alternative care, charges for phone contacts with providers, or charges that have not been paid are not included in the calculations of expenditures.

In addition, follow-up expenditure data on the reported utilization events is collected for most utilization categories from the health care providers and pharmacies that the participants reported using. Expenditure information reported by providers is considered more accurate by the MEPS and is used as the primary source
of expenditure data (except where it is not collected such as for dental services), while participant reported data are used for utilization event count.

For each person, the utilization and expenditure information collected is summed over the year across these event type categories. For each event category, the MEPS provides a count of utilization events, total charges for the category, and 13 expenditure variables, one with total payments made and 12 for specific sources of payment. In addition, an aggregate total annual expenditure value is provided, which sums the expenditures over approximately 2.5 interviews per household across all events. Total expenditure amounts for the year were used as an outcome in paper 3. These were converted into 2010 dollars using the Bureau of Labor Statistics consumer price index.  

**c. Chronic Health Conditions**

Starting in year 2000, questions about prior diagnosis of several medical conditions were included in the MEPS survey rounds 3 and 5, and the conditions included were different from those included in the NHIS. The conditions were diabetes, asthma, high blood pressure, coronary heart disease, angina, myocardial infarction, any other heart disease, stroke, and emphysema. The participants were asked if they have ever been diagnosed with these conditions by a health care professional. In 2001, a question about lifetime diagnosis of arthritis was added. As arthritis diagnosis is an important predictor of HRQL and health service utilization, the chronic condition variable created for papers 2 and 3 included only the diagnosis data starting from year 2001. A heart disease (yes/no) variable was
created for this research that was coded as “yes” if either one of the following has been previously diagnosed: coronary heart disease, angina, myocardial infarction, or any other heart disease. The chronic condition variable used for modeling the outcomes was a count of the following prior diagnoses: diabetes, asthma, high blood pressure, heart disease, stroke, emphysema, and arthritis.

d. Smoking

Current smoking question (yes/no) became a part of SAQ in 2000. Prior history of smoking was not assessed in the MEPS, and alcohol consumption patterns were not assessed either.

e. Education

Education was represented as the number of years of education at the time the participant entered the MEPS and was treated as a continuous variable.

f. Age

Age was treated as a continuous variable. Until year 2000, age was top-coded in the MEPS at 90 years old, however starting in 2001 it was top-coded at 85. All the age values greater than 85 in this study’s dataset were recoded to 85 for consistency (n=1,245).
g. **Race/ethnicity**

Race and Hispanic ethnicity were assessed for every participant at least once during the MEPS survey. If the information could not be obtained from the interview, it was obtained from the original NHIS data or derived from the races/ethnicities of the participant’s blood relatives. Two variables, race and Hispanic ethnicity were used to create the following categories for this dissertation: 1) non-Hispanic white; 2) non-Hispanic black; 3) Hispanic; and 4) other.

h. **Gender**

Gender information was derived from the NHIS data, and was verified and corrected if needed during each MEPS interview. This was coded as 1) male, 2) female.

i. **Employment / Occupation**

Employment and occupation were assessed during each round of the MEPS. During each round participants were asked to identify the main job, which was included in the full-year consolidated file, however other jobs were recorded as well in the jobs file. To assess the employment status, several questions were asked about whether the person currently had a job, or owned a business, or had a job to return to, and if not whether they had worked at any time during the round reference period. For these analyses, those who responded with options “currently employed”, “have a job to return to,” and “employed during the round reference period”, were classified as employed.
The occupation information recorded verbatim was recorded by the trained Bureau of Census coders into a condensed classification based on 1990 and 2000 SOC codes. Prior to 2002, this was condensed to 8 categories, and from 2002 and on this was 11 categories. For this research, these categories were combined with employment information and recoded into: 1) unemployed/retired; 2) white collar, 3) blue collar, 4) farm worker, and 5) service worker. No distinction was made between unemployed and retired individuals.

Employment/occupation information utilized for HRQL analyses was measured at the same time as HRQL scores (rounds 2 and 4). For expenditures and health care utilization modeling, this research used the employment/occupation information measured at the beginning of the year of interest (rounds 1 and 3).

Previous retirement status was measured in MEPS for individuals who reported ever previously having had a job. At the end of each survey year (rounds 3 and 5) these participants were asked if they ever retired from a job. Approximately 16% of the participants, all of which reported being currently unemployed, did not have retirement information available. Of those who answered this question, 67.1% reported having previously retired (Table 2.1).

<table>
<thead>
<tr>
<th>Employment / Occupation</th>
<th>Males, % (SE of %)</th>
<th>Females, % (SE of %)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Never</td>
<td>Previously</td>
</tr>
</tbody>
</table>
With employment/occupation variables in our study being measured during rounds 1 & 3 or 2 & 4, retirement status was assessed several months to almost a year after employment variables of interest in our study were measured, which at post-retirement age is a long time for employment status to change. Therefore combining employment/occupation and retirement status information available in one model was deemed inappropriate, and it was left out of the analyses. Retirement status information was also not included given: 1) the large number of missing values for this variable, 2) the creation of very small sample sizes within some occupation*retirement sub-groups (e.g., farmers), and 3) the fact that all of the missing fell into the currently unemployed category, thereby introducing the strong possibility of rendering the sample non-representative of the older US population.

\[ j. \quad \textit{Health Insurance Coverage} \]

Health insurance coverage was assessed for every month of the year and included coverage under Tricare, Medicare, Medicaid, other public hospital/physician insurance, or private hospital/physician insurance (including Medigap plans).

<table>
<thead>
<tr>
<th>Round 2 or 4</th>
<th>worked</th>
<th>retired</th>
<th>retired</th>
<th>worked</th>
<th>retired</th>
<th>retired</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unemployed</td>
<td>6.39</td>
<td>81.17 (0.67)</td>
<td>12.44 (0.52)</td>
<td>21.96</td>
<td>47.26 (0.72)</td>
<td>30.79 (0.71)</td>
</tr>
<tr>
<td>White Collar</td>
<td>NA</td>
<td>50.21 (1.95)</td>
<td>49.79 (1.95)</td>
<td>NA</td>
<td>42.11 (1.65)</td>
<td>57.89 (1.65)</td>
</tr>
<tr>
<td>Blue Collar</td>
<td>NA</td>
<td>52.08 (2.55)</td>
<td>47.92 (2.55)</td>
<td>NA</td>
<td>35.51 (6.35)</td>
<td>64.49 (6.35)</td>
</tr>
<tr>
<td>Farmer</td>
<td>NA</td>
<td>45.91 (8.05)</td>
<td>54.09 (8.05)</td>
<td>NA</td>
<td>14.61 (9.50)</td>
<td>85.39 (9.50)</td>
</tr>
<tr>
<td>Service</td>
<td>NA</td>
<td>63.92 (3.56)</td>
<td>36.08 (3.56)</td>
<td>NA</td>
<td>37.95 (2.90)</td>
<td>62.05 (2.90)</td>
</tr>
</tbody>
</table>
Individuals whose insurance did not cover hospital and physician service coverage were classified as “uninsured.” This included state programs or private insurance plans that did not provide comprehensive coverage (e.g. Maryland Kidney Disease Program, plans providing dental or drug coverage only, etc.). Aggregate variables were created in the MEPS that indicated coverage by a specific type of insurance (public, private), specific insurance source (Medicare, employer, Tricare, etc.), or coverage at a specific point in time or during the entire year.

The MEPS collects a variety of information on insurance offered to or held by the participant through employer. In addition, information on public sources of insurance in any month is collected. The public sources of insurance are Medicare, Tricare, Medicaid and other public hospital/physician coverage. Private insurance coverage was recorded as such if it provided at least hospital. Private coverage included Medigap.

A variable summarizing insurance coverage over the year was included for analyses in the modeling outcomes using the MEPS data; it was coded as: 1) any private insurance coverage during the year (this variable included TRICARE /VA under private); 2) public only during the year; and 3) uninsured during the entire year (including those covered by a single-service plan such as vision, dental, drug). For this variable, Tricare was categorized as private coverage. For all other insurance variables in the MEPS, Tricare is categorized as public insurance. Even though a greater proportion of the participants reported having private insurance coverage, the public category was treated as the reference group since it is the source for a greater proportion of older adult healthcare spending (Figure 2.6).
Figure 2.6: Sources of Payment for Healthcare Expenses among Medicare Beneficiaries Aged 65+. 2007 Medical Expenditure Panel Survey.


**k. Income**

The income variable was coded based on the family income as a percentage of poverty level. This included all sources of income except for tax refunds and capital gains. Personal income amounts were summed for all family members and divided by the current poverty level for the given family size. This variable was classified as: 1) poor (less than 100% of poverty level); 2) near poor (100%-125%); 3) low income (125%-200%); 4) middle income (200%-400%); 5) high income (≥400% of poverty level).
VI. Statistical Analysis Methods

General considerations for statistical analyses are described first in this section. Further, details of the analyses are organized by Specific Aim / Paper.

i. Data Aggregation and Complex Survey Design Effects

All estimates, standard errors, and test statistics were adjusted for complex sample design effects using clustering, stratification, and sampling weights information provided in the data and using robust estimators in SAS and Mplus. Sampling weights provided in the NHIS and MEPS datasets were usually inappropriate when data are aggregated across survey years. Therefore weights were adjusted by dividing them by the number of years the data was pooled for, in accordance with the work of Botman and Jack. Thus, the weights used with the NHIS data were divided by 15, the weights used for modeling SF-12 components were divided by 10, and the weights used for modeling EQ-5D components were divided by 4.

ii. Regression Modeling

The distribution of outcomes variables was examined to determine which type of regression modeling would be most appropriate. Logistic regression modeling was utilized for binary outcomes; linear regression modeling for normally distributed continuous and count outcomes; and Poisson and negative binomial regression modeling for count outcomes with a non-normal (Poisson) distribution. All of these are referred to comprehensively as “regression modeling” in this chapter.
iii. Statistical Packages

SAS Version 9.3 (SAS Institute, Cary, NC) was used for all data management (i.e. data cleaning, concatenating, merging, variable coding) and descriptive analyses. SAS and Mplus Version 6.1 (Muthén & Muthén, Los Angeles, CA) were used for analytical modeling because these packages are able to apply the appropriate weights and adjust for design effects.

iv. Structural Equation Modeling (SEM)

SEM is a general modeling framework that incorporates many common statistical methods including generalized linear regression, ANOVA, confirmatory factor analyses, and simultaneous equations.\textsuperscript{128,129} SEM offers several advantages over traditional analytic methodology. Of specific interest for this research is that SEM allows for the estimation of multiple equations simultaneously so that associations between multiple predictor and outcome variables can be assessed within the same model. In this study, SEM provided a powerful tool for the simultaneous assessment of complex relationships between multiple variables including mediation (see Conceptual Model, Figure 1.1)\textsuperscript{130}.

SEM modeling was carried out using Mplus. For each of the outcomes, structural equation path models were created, in which socio-demographic and health behavior variables were placed as exogenous variables, and health insurance, income, number of health conditions, and employment/occupation were placed as mediators. At the initial stage, all possible paths were included in the models, and following each
iteration, insignificant paths were eliminated. Therefore, not all paths in Figures 2.1 and 2.2 are present in each of the final models.

v. Missing Data

Listwise deletion was utilized for treatment of missing data in the NHIS-based analyses. Listwise deletion is the default approach to handling missing data in SAS, and it results in observations being excluded from the analyses if information is missing on at least one of the variables included in the model.

In contrast, Mplus uses pairwise deletion with a full information maximum likelihood (FIML) approach to treatment of missing data. FIML requires that the data are missing in part at random (MAR), or missing completely at random (MCAR). FIML estimates a likelihood function for each individual observation based on the data that are available on that observation, thereby allowing the use of all available data.

Maximum likelihood with robust standard errors (MLR) estimator was used for modeling in Mplus. MLR standard errors are robust to non-normality and non-independence of complex survey data, and FIML for the treatment of missing data.

vi. Analytic Approach to the Specific Aims

The analytic approach varied by specific aim depending on its objectives. The description below is a summary of the methodologies used for each specific aim; additional information can be found in the corresponding chapters.
Chapter 3 presents the results of **Specific Aim 1** analyses. Descriptive analyses were conducted examining odds of multimorbidity, multiple functional limitations, fair/poor self-rated health, and HALex HRQL measure among subgroups of US older adults, including workers and non-workers. Descriptive statistics (including means and 95% confidence intervals for the number of chronic conditions and HALex were presented, and percents and 95% confidence intervals for discrete data) were presented for variables of interest for the overall sample as well as by gender.

Using multivariable logistic regression, the odds of fair/poor self-rated health, multiple functional limitation, multiple chronic conditions, and HALex among older adult subgroups, such as workers of different occupations, were evaluated while controlling for sociodemographic factors and health behaviors, such as smoking status, race/ethnicity, gender, etc. Further description of this Aim’s methodology is included in chapter three.

Chapter 4 presents the results of **Specific Aim 2** analyses. Using SEM and the SF-12/SF-12v2 and EQ-5D HRQL data available in the 2003-2009 MEPS, the quality of life of older adult subpopulations was explored. SEM techniques were used for the prediction of quality of life in different older worker and non-worker subpopulations after adjustment for health behaviors and co-morbidities. Mediation effects of chronic conditions, employment / occupation status, income, and health insurance coverage were tested. The hypotheses tested were the following:

\[ H_{2.1}: \text{While workers in white collar occupations who continue working past retirement age will demonstrate higher quality of life relative to retired peers, work past retirement age for workers in other occupations will result in lower} \]
quality of life relative to their retired peers after adjustment for other factors (MEPS).

H2.2: While older workers in race/ethnic minority subgroups will experience lower quality of life than non-Hispanic white older workers, this effect will be eliminated by controlling for occupational group (MEPS).

SEM was used to estimate models based on the Conceptual Model in Figure 2.1, which was based on a more complex model in Figure 1.1. SF-12v2 and EQ-5D scores were treated as observed variables in an SEM model. Estimates from different pathways varied based on the functional form of the outcomes. For example, Mplus with MLR estimator uses linear regression for modeling continuous outcomes; probit or logistic regression for modeling binary and ordinal outcomes; multinomial logistic regression for unordered categorical outcomes; and Poisson or negative binomial regression for count outcomes. Therefore, within the same model, some pathways were estimated using ordinary linear regression parameters, while other pathways with non-continuous outcomes were estimated using probit, or logit, parameters.

To model categorical mediators (employment/occupation and health insurance status) in Mplus, these variables had to be recoded as dummy variables, with a separate dummy variable representing each of the categories of the original variables. Because of the presence of categorical mediators in the model, standardized path coefficients could not be obtained using MLR. Mediated pathways were estimated and tested using a product of coefficients method and a WLS estimator (treating categorical variables as continuous, as indirect path estimates could not be obtained with MLR).
The model $\chi^2$ was statistically significant for all models due to the large sample size. Therefore, alternative methods had to be used to estimate model fit. The Root Mean Square Error of Approximation (RMSEA) and the Comparative Fit Index (CFA), the standard fit indices used to evaluate SEM model fit, could not be obtained with the MLR estimator. Therefore model fit was assessed using the Satorra-Bentler scaled chi-square difference test. The $\chi^2$ for the test is calculated as follows:

$$cd = (p_0 \times c_0 - p_1 \times c_1)/(p_0 - p_1)$$

$$TR_d = -2(L_0 - L_1)/cd$$

where cd is the difference test scaling correction, p0 is the number of parameters in the nested model, p1 is the number of parameters in the comparison model, c0 and c1 are the scaling correction factors for the nested and the comparison models respectively, and L0 and L1 are the log likelihood values for the nested and the comparison models, respectively. The resulting value, TRd, can be used for chi-square difference testing of nested models with MLR estimator.

Chapter 5 presents the results of Specific Aim 3 analyses. Using SEM, the health care utilization and expenditures were compared in working and non-working US adults 65+. These outcomes were controlled for health insurance coverage (public/private/none) and the number of co-morbidities. The Conceptual Model in Figure 2.2, which was based on the general conceptual model in Figure 1.1.

$H_{3.1}$: Total healthcare expenditures will be significantly lower in older workers versus older non-workers, even after adjustment for sociodemographic factors and co-morbidities;
**H3.2:** Total healthcare expenditures will be significantly higher in blue collar than in white collar older workers, and this relationship will be only partially mediated by the type of health insurance coverage.

SEM techniques were used as described above (see Specific Aim 2 Analytic Approach description above) to predict healthcare utilization (number of annual ER visits and nights of hospital stay) and total expenditures in various older worker subpopulations based on the Conceptual Model in Figure 2.2. All outcomes were count variables and were highly skewed. Therefore, Poisson regression modeling was used to model ER visits, and due to a great degree of dispersion in total expenditures and the nights of hospital stay, these outcomes were modeled using negative binomial regression.

Due to a large number of zero values for all outcomes, the analyses had to be zero-inflated. In zero-inflated Poisson or negative binomial regression, modeling consists of two parts which are estimated simultaneously. The first part of the regression evaluates which factors predict having a non-zero outcome. Next, a Poisson or negative binomial regression evaluates which factors are associated with higher values of outcomes in those with non-zero values. The two parts do not have to use the same predictors.

The Poisson/negative binomial part of the model estimates the log of change in the outcome measured. Therefore an incidence rate ratio (IRR) for outcome y due to a one unit change in predictor x can was calculated with the path coefficients provided by Mplus output using the following formula: $\text{IRR}_y = (e^{\beta_x})$, where $\beta_x$ is the path coefficient estimate for predictor x. For expenditures, mean expenditures for
each group were calculated after adjusting for the probability of a non-zero expenditure.

VII. Human Subjects Protection

This research was reviewed by the University of Miami Institutional Review Board and ruled as “exempt” (Study Number #20110677) i.e. research involving the study of existing data. It used de-identified data that is publically available for download and use.
Chapter III: Specific Aim / Paper #1

I. Background

Adults aged 65+ are a rapidly expanding segment of the US population, and a growing proportion of them are active in the workforce. Characterization of the health status of this population has been complicated by the lack of consensus about most appropriate and relevant measures. The most commonly used measures of older adult health status have included presence of functional limitations or disability, self-reported health or health-related quality of life (HRQL), and chronic health conditions. Studies using nationally-representative data have found that between 26.0% and 33.2% of adults aged 65+ currently report being in fair/poor health compared to 7.3-14.4% among adults aged less than 55, and that approximately 27.9-48.3% of adults aged 65+ report great difficulty in at least one of the activities assessing daily function. Estimates of multimorbidity prevalence in this population have ranged from 47% to 73%. Interpretation of these results is often complicated by the lack of consensus on the definitions of such outcomes as disability, functional limitation, or HRQL.

Regardless of the measure used, a great deal of variation in the health status is evident across socio-demographic groups of older adults. For example, HRQL scores vary across income, education levels, gender, and racial/ethnic groups. Blacks are more likely than whites to report a functional limitation even after adjustment for age and gender. Women, racial/ethnic minorities, and individuals of low socio-economic status are especially likely to report having a disability. In addition, employed older individuals tend to be healthier, both mentally and physically, than their non-working peers.
The above studies were subject to some major limitations including reporting only the prevalence of the health outcomes without controlling for potential confounders, and using data not representative of the entire US population. As a result, these may have led to incomplete conclusions. In the current study, we characterize four major health status measures of older US adults using a representative sample of US population and controlling for potential socio-demographic and health behavior confounders such as education, race/ethnicity, gender, age, smoking/drinking habits, and survey year. The results of this study will allow identification of disparities within the heterogeneous aging population to assist in older adults’ workplace need assessment.

II. Methods

i. Sample

The National Health Interview Survey (NHIS) is an annual multistage probability household survey of the US civilian non-institutionalized population. It uses face-to-face interviews to obtain socio-demographic and health characteristics information. This study included a sample of all adults aged 65 and older (n=83,338; representing approximately 33,546,235 individuals) from the NHIS cross-sectional data pooled over years 1997-2011. Sampling weights adjustment were used to ensure unbiased estimates of the national population. This study was approved by the University of Miami Institutional Review Board.
ii. Variables

The four main outcomes examined in this study were: self-rated health, multimorbidity, functional limitations, and Health and Activities Limitation Index (HALex, a measure of HRQL). The predictors included in all models were employment/occupation, education, race/ethnicity, gender, age, smoking, and alcohol consumption. Analyses were adjusted for the survey year in order to control for potential change in prevalence of the outcomes over time. Outcome variables were dichotomized for easier comparison with previous studies.

**Self-rated health.** Participants rated their perceived health on a 5-point Likert scale: 1) excellent, 2) very good, 3) good, 4) fair, 5) poor. These were collapsed into 1) fair/poor and 2) good or better categories, which was shown to yield results similar to the multiple-category variable.\(^{144}\)

**Multimorbidity.** Multimorbidity was defined as previous lifetime diagnosis of two or more of the following conditions: hypertension, heart disease (including coronary heart disease, angina, and myocardial infarction), stroke, emphysema, asthma, cancer, and diabetes.\(^{110}\)

**Multiple functional limitations.** The number of functional limitations was assessed with the question, “By yourself, and without using any special equipment, how difficult is it for you to...” with regards to the following activities: walking, climbing stairs, standing, sitting, stooping or kneeling, reaching over head, grasping, carrying heavy objects, pushing large objects, shopping, being social, and relaxing. Participants rated their ability to perform these activities on a five-point scale from ‘not at all difficult’ to ‘can’t do at all’. A functional limitation was defined as having
any difficulty with any one of the above activities. The presence of multiple functional limitations was defined as having limitations with 2 or more of these activities.

**HALex.** HALex index was calculated as described by Livingston and Ko.\textsuperscript{113} HALex is a utility score combining information about an individual’s perceived health and activity limitations that ranges from 0 to 1, with 1 representing perfect health, and 0 representing a health state equivalent to death. The activity limitations included needing help with personal care or routine needs, having difficulty working, being limited in the kind and amount of work, or being limited in any other way due to health reasons. Participants were dichotomized into those below and above the 20th percentile value for the population represented (i.e. HALex≤0.48 and HALex>0.48). The 20\textsuperscript{th} percentile was chosen as the cutoff value in order to model the lowest HALex scores while allowing for sufficient sample size in all population subgroups examined.

**iii. Predictor variables.**

The primary predictor variable in this study is an employment / occupation type hybrid variable. The employment/occupation variable combined information about whether the individual had worked in the week prior to NHIS interview (full-time or part-time) and the kind of work they reported doing. It was classified as 1) unemployed/retired [reference], 2) white-collar workers, 3) service workers, 4) farm workers, and 5) blue-collar workers.\textsuperscript{109} No distinction was made between temporarily unemployed and retired individuals.
The effect of the following variables on the outcomes was examined as well: education (less than high school [reference], high school or equivalent, more than high school), race/ethnicity (non-Hispanic white [reference], non-Hispanic black, Hispanic, other), gender (male, female [reference]), age (continuous), smoking history (current, former, never [reference] smoker), and alcohol consumption history (current heavy, current light, former, never [reference] drinker).

iv. Statistical Analysis

Multiple logistic regression analyses were used to test the associations between employment/occupation and the health outcomes controlling for covariates. Data management and analyses were performed using SAS version 9.3 (SAS Institute Inc., Cary, NC). Sample Adult NHIS file sampling weights were used to correct for the unequal selection design of the NHIS. In addition, standard errors were adjusted for the nesting of persons within sampling clusters and the added variance of the sampling weights. Models were tested for the possible interactions between gender and occupation to control for the potential differential effects of employment/occupation on the health outcomes for male and female older workers.6

III. Results

The population had a mean age of 74.6 years (Table 3.1). Over a half were female (57.2%), and the majority were non-Hispanic white (82.5%), unemployed/retired (87.1%), and reported no history of heavy drinking in the previous year (95.6%). Approximately two thirds of those employed worked in white-collar professions (7.3% of
Table 3.1: Sample characteristics. National Health Interview Survey (NHIS) 1997-2011 participants aged 65+.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>All</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NHIS Frequency</td>
<td>Weighted % (SE of %)</td>
<td>NHIS Frequency</td>
</tr>
<tr>
<td>Total</td>
<td>83,338</td>
<td>100</td>
<td>31,988</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than high school</td>
<td>25,979</td>
<td>27.5 (0.27)</td>
<td>9,688</td>
</tr>
<tr>
<td>High school or equivalent</td>
<td>26,710</td>
<td>33.2 (0.23)</td>
<td>8,965</td>
</tr>
<tr>
<td>Greater than high school</td>
<td>30,649</td>
<td>39.3 (0.29)</td>
<td>13,335</td>
</tr>
<tr>
<td>Race/Ethnicity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>7,778</td>
<td>6.3 (0.16)</td>
<td>3,048</td>
</tr>
<tr>
<td>Non-Hispanic white</td>
<td>63,025</td>
<td>82.5 (0.27)</td>
<td>24,283</td>
</tr>
<tr>
<td>Black</td>
<td>9,933</td>
<td>8.2 (0.19)</td>
<td>3,572</td>
</tr>
<tr>
<td>Other</td>
<td>2,602</td>
<td>3.1 (0.12)</td>
<td>1,085</td>
</tr>
<tr>
<td>Smoking Status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>43,331</td>
<td>50.4 (0.25)</td>
<td>11,121</td>
</tr>
<tr>
<td>Former</td>
<td>31,554</td>
<td>40.0 (0.24)</td>
<td>17,170</td>
</tr>
<tr>
<td>Current</td>
<td>8,453</td>
<td>9.6 (0.13)</td>
<td>3,697</td>
</tr>
<tr>
<td>Drinking Status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>27,059</td>
<td>30.1 (0.29)</td>
<td>5,590</td>
</tr>
<tr>
<td>Former</td>
<td>21,905</td>
<td>25.9 (0.21)</td>
<td>9,896</td>
</tr>
<tr>
<td>Current light</td>
<td>30,983</td>
<td>39.6 (0.28)</td>
<td>13,861</td>
</tr>
<tr>
<td>Current heavy</td>
<td>3,391</td>
<td>4.4 (0.09)</td>
<td>2,641</td>
</tr>
<tr>
<td>Employment/Occupation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unemployed/retired</td>
<td>73,138</td>
<td>87.1 (0.15)</td>
<td>26,942</td>
</tr>
<tr>
<td>White collar</td>
<td>6,115</td>
<td>7.9 (0.12)</td>
<td>2,693</td>
</tr>
<tr>
<td>Service</td>
<td>2,057</td>
<td>2.4 (0.06)</td>
<td>737</td>
</tr>
<tr>
<td>Farmer</td>
<td>281</td>
<td>0.3 (0.03)</td>
<td>230</td>
</tr>
<tr>
<td>Blue collar</td>
<td>1,747</td>
<td>2.3 (0.07)</td>
<td>1,386</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Weighted Mean(SE)</th>
<th>95% CI</th>
<th>Weighted Mean(SE)</th>
<th>95% CI</th>
<th>Weighted Mean(SE)</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>HALex</td>
<td>0.74 (0.00)</td>
<td>0.74-0.74</td>
<td>0.75 (0.00)</td>
<td>0.75-0.75</td>
<td>0.73 (0.00)</td>
<td>0.72-0.73</td>
</tr>
<tr>
<td>Age</td>
<td>74.59 (0.04)</td>
<td>74.52-74.67</td>
<td>73.99 (0.05)</td>
<td>73.89-74.08</td>
<td>75.05 (0.05)</td>
<td>74.96-75.14</td>
</tr>
<tr>
<td>Number of comorbidities</td>
<td>1.40 (0.01)</td>
<td>1.39-1.41</td>
<td>1.48 (0.01)</td>
<td>1.47-1.50</td>
<td>1.34 (0.01)</td>
<td>1.33-1.35</td>
</tr>
</tbody>
</table>

aThe count of chronic conditions derived from NHIS data included: hypertension, stroke, emphysema, asthma, cancer, heart disease, and diabetes.
the total sample), and only a small fraction (0.3% of the sample) were farm workers. Approximately half of the population (50.4%) reported never smoking, with 9.6% currently smoking. Descriptively, women were less likely to report poor health behaviors: 62.0% never smoked and only 1.6% reported current heavy drinking as compared to 35.0% and 8.2% for males, respectively. Women were also less likely to be educated beyond high school (35.4% vs. 44.6%), and more likely to not work (89.9% vs. 83.3%) compared to males.

Table 3.2: Multiple Logistic Regression Results for the Health Indicators. All National Health Interview Survey 1997-2011 Participants Aged 65+

<table>
<thead>
<tr>
<th></th>
<th>Fair/Poor Health</th>
<th>Multiple Functional Limitations</th>
<th>Poorest HALex Quintile</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Odds Ratio(95% CI)</td>
<td>Odds Ratio(95% CI)</td>
<td>Odds Ratio(95% CI)</td>
</tr>
<tr>
<td>Education (vs. Less than High School)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High School</td>
<td>0.60 (0.57-0.63)</td>
<td>0.71 (0.67-0.74)</td>
<td>0.61 (0.58-0.64)</td>
</tr>
<tr>
<td>More than High School</td>
<td>0.43 (0.41-0.46)</td>
<td>0.62 (0.59-0.65)</td>
<td>0.52 (0.49-0.56)</td>
</tr>
<tr>
<td>Race/Ethnicity (vs. White non-Hispanic)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>1.64 (1.54-1.75)</td>
<td>0.85 (0.79-0.90)</td>
<td>1.25 (1.16-1.33)</td>
</tr>
<tr>
<td>Black</td>
<td>1.68 (1.58-1.79)</td>
<td>1.08 (1.02-1.14)</td>
<td>1.53 (1.43-1.64)</td>
</tr>
<tr>
<td>Other</td>
<td>1.28 (1.16-1.41)</td>
<td>0.82 (0.74-0.90)</td>
<td>1.06 (0.94-1.21)</td>
</tr>
<tr>
<td>Gender (vs. Female)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>1.15 (1.10-1.20)</td>
<td>0.61 (0.58-0.63)</td>
<td>0.83 (0.78-0.87)</td>
</tr>
<tr>
<td>Age</td>
<td>1.02 (1.01-1.02)</td>
<td>1.05 (1.05-1.06)</td>
<td>1.05 (1.05-1.06)</td>
</tr>
<tr>
<td>Smoking (vs. Never)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Former</td>
<td>1.37 (1.31-1.44)</td>
<td>1.29 (1.25-1.35)</td>
<td>1.41 (1.34-1.48)</td>
</tr>
<tr>
<td>Current</td>
<td>1.63 (1.53-1.75)</td>
<td>1.43 (1.35-1.52)</td>
<td>1.84 (1.70-1.99)</td>
</tr>
<tr>
<td>Alcohol (vs. Never)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Former</td>
<td>1.03 (0.98-1.09)</td>
<td>1.28 (1.22-1.34)</td>
<td>1.12 (1.06-1.18)</td>
</tr>
<tr>
<td>Current light</td>
<td>0.47 (0.44-0.49)</td>
<td>0.74 (0.71-0.78)</td>
<td>0.45 (0.42-0.47)</td>
</tr>
<tr>
<td>Current heavy</td>
<td>0.43 (0.38-0.48)</td>
<td>0.88 (0.81-0.96)</td>
<td>0.41 (0.35-0.48)</td>
</tr>
<tr>
<td>Employment/Occupation (vs. Unemployed/retired)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White-collar worker</td>
<td>0.36 (0.33-0.41)</td>
<td>0.54 (0.50-0.58)</td>
<td>0.17 (0.14-0.21)</td>
</tr>
<tr>
<td>Service worker</td>
<td>0.39 (0.33-0.47)</td>
<td>0.44 (0.39-0.50)</td>
<td>0.14 (0.11-0.19)</td>
</tr>
<tr>
<td>Farm worker</td>
<td>0.30 (0.20-0.45)</td>
<td>0.52 (0.39-0.69)</td>
<td>0.16 (0.08-0.35)</td>
</tr>
<tr>
<td>Blue-collar worker</td>
<td>0.42 (0.36-0.49)</td>
<td>0.46 (0.40-0.53)</td>
<td>0.15 (0.11-0.21)</td>
</tr>
<tr>
<td>Survey year</td>
<td>1.00 (1.00-1.01)</td>
<td>1.01 (1.01-1.02)</td>
<td>1.01 (1.00-1.01)</td>
</tr>
</tbody>
</table>
Multivariable logistic regression results are presented in Table 3.2 (modeling multiple functional limitations, poor/fair self-rated health, and poorest HALex) and Table 3.3 (modeling multimorbidity stratified by gender based on significant gender-interaction results). Employment/occupation was strongly associated with all outcomes. Employed older individuals of all occupations had better health as measured by all outcomes when compared to their non-working peers, with service workers
demonstrating the best outcomes followed by farm and blue-collar workers. Service
workers were at the lowest risk of multiple functional limitations (Odds Ratio=0.44; 95%
Confidence Interval 0.39-0.50) and of poorest HALex quintile (0.14; 0.11-0.19). Farm
workers had the lowest risk of fair/poor health (0.30; 0.20-0.45), and blue-collar workers
had the lowest risk of multimorbidity (0.54; 0.48-0.62). In addition, farm workers had the
second lowest risk of multimorbidity (0.60; 0.44-0.83), while blue-collar workers had the
second lowest risk of multiple functional limitations (0.46; 0.40-0.53) and of being in the
lowest HALex quintile (0.15; 0.11-0.21). Employment/occupation significantly interacted
with gender in predicting multimorbidity (p<0.0001), therefore the results for this
outcome were stratified by gender. For males, blue-collar (0.52; 0.45-0.61) and farm
workers (0.64; 0.45-0.91) were at the lowest risk for multimorbidity; for females the
lowest risk was among white-collar (0.59; 0.53-0.65) and service workers (0.59; 0.51-
0.69).

Education, race/ethnicity, gender, age, and health behaviors were all associated
with health status measures. Non-Hispanic blacks had worse health outcomes than non-
Hispanic whites. Hispanics were at a higher risk than non-Hispanic whites for reporting
fair/poor health and for being in the lowest HALex quintile, however they were less
likely to report multimorbidity and functional limitations. Males were more likely than
females to report fair/poor health and multimorbidity, however they were less likely to
report multiple functional limitations and be in the lowest HALex quintile. Compared to
less than high school education, higher levels of education were associated with better
health outcomes. Compared to never consuming alcohol, former alcohol consumption
was associated with poorer health outcomes, while current alcohol consumption was
associated with better health outcomes even for heavy drinkers. Both former and current smoking was associated with poorer health outcomes across all outcomes when compared to never smoking.

IV. Discussion

In this nationally representative sample of adults aged 65+, the typical older individual was a non-smoker aged approximately 74 years old with one or two chronic health conditions who did not work. The majority of employed individuals worked in white-collar occupations, only a small fraction worked in farming, and the rest were approximately equally distributed between blue-collar and service occupations. Employment in any occupation as compared to being unemployed/retired was associated with the greatest reduction in the risk of poor health across all health status measures. Older adults in more physically demanding occupations (such as service, farming, and blue-collar) seemed to fare better than those in white-collar occupations. This is possibly due to a stronger healthy worker effect among these occupations, as the healthier individuals were more likely to continue working, while those in poorer health were more likely to exit the workforce.\(^ {145}\) However, working at older age can also have a beneficial effect on health by increasing one’s social involvement, social support, and physical activity levels, and by providing access to more comprehensive health insurance coverage, thereby acting to improve the health status of those older adults who continue working.\(^ {12,13,50}\)

Consistent with previous studies, gender, education, race/ethnicity, age, and drinking/smoking history were associated with health status measures.\(^ {135,142,146,147}\) Poorer
outcomes across all measures were associated with less than high school education as compared to any higher level of education, as well as with being non-Hispanic black as compared to non-Hispanic white race/ethnicity. Former alcohol consumption was associated with a greater risk of all poor outcomes except fair/poor self-rated health, while current drinking was associated with better health outcomes, even for heavy drinkers. Alcohol abstinence in older adults has been linked with loss of mobility and a higher risk of dementia,\textsuperscript{148,149} while low to moderate alcohol consumption was associated with lower overall mortality\textsuperscript{150} and higher health-related quality of life,\textsuperscript{151} and heavy alcohol consumption was associated with higher bone mineral density.\textsuperscript{152} These possibly indicate a genetic resiliency in older drinkers, with those most susceptible to the negative effects of alcohol either not surviving to older age or quitting drinking at a younger age.

While most predictors resulted in consistent associations across outcomes, some had different effects depending on the outcome modeled. Hispanics were at a higher risk for being in fair/poor health and poorest HALex quintile; however they were less likely to report multimorbidity and functional limitations. Poorer outcomes on health status measures incorporating self-perceived health in older Hispanics might be partially due to a difference in perception across cultures of what constitutes good health, as well as to limited access to health care in this group.\textsuperscript{124,153} Issues with access to care are also suggested by the lower risk of multimorbidity in Hispanics, possibly a result of chronic conditions being under-diagnosed in this group.\textsuperscript{64,154} However, limited access to care does not explain the lower risk of multiple functional limitations in this group that is consistent with previous frailty studies,\textsuperscript{155} and this discrepancy should be examined in future studies.
We also found that males were more likely than females to report fair/poor health and multimorbidity; however, males were less likely to report multiple functional limitations and to be in the lowest quintile for the HALex scores. That is, women did worse on both measures that were based on functional limitations, while they were more likely to perceive their health as good or excellent and reported fewer health conditions. In previous studies, individuals in poor health did not necessarily report disability or limitations, and functional limitations were more likely to be reported if a person had limited availability and access to assistive devices.\textsuperscript{137,156,157} In addition, some of the observed differences might be due to underreporting of functional limitations by men.\textsuperscript{158} Finally, the multimorbidity variable in our study did not include such disabling conditions as osteoporosis and resulting fractures, which are more common in women than men, possibly resulting in increased reports of limitations by women without increasing their risk of multimorbidity in our results.\textsuperscript{159}

V. Strengths and Limitations

This study used pooled cross-sectional data, and therefore causal inferences cannot be made. Past employment history information was not available, and while this study aimed to examine the effects of employment at older age, such effects may vary depending on whether the person re-entered the workforce after retirement, and whether they changed jobs or remained in career employment. Inability to work due to health reasons was one of the components used in calculation of Halex, and this might have led to over-inflation of the association between employment and Halex, albeit to a small degree. While a moderately high agreement was demonstrated between longest held job
and current job with past NHIS data, this might not apply to older workers who are more likely to retire from their life-time career employment and then seek new less demanding employment.

The major strength of this study was a large nationally representative dataset obtained by pooling 15 years of data, with information on a range of socio-demographic and health status variables. The multivariable regression analysis utilized in this study allowed examining the effects of multiple factors while controlling for the effects of potential confounders, and it was an improvement on previous prevalence reporting studies. We also used four different but complementary health status measures in order to most comprehensively characterize the health of this population.

VI. Conclusions

The prevalence of chronic conditions increases as people live longer with diseases, however this does not necessarily translate into increased prevalence of functional limitations. These various aspects of individual’s health not only affect the individual’s functioning differently, but also require different amounts and kinds of health care resources in order to address different needs of the population. In addition, as a growing number of older adults stay active in the workforce, occupational health resources need to be allocated differently to address the needs of the aging workforce. Older adults who continue working tend to be much healthier across multiple health outcomes, but perhaps providing better workplace accommodations for older adults with functional limitations would allow more of them to join the ranks of their healthier peers. Determination of the factors most closely associated with poor health outcomes across
various health measures is therefore important for fine-tuning healthcare resource allocation, both in and outside the workplace, as the population ages.

In the current study, we characterized the health of the older US workers and non-workers by examining the risk of four complimentary poor health outcomes across various socio-demographic groups. We found a strong association between health status and employment/occupation, as well weaker associations for the following: education, race/ethnicity, gender, and smoking/drinking history. We also identified a variation in health status across different measures within the same population subgroups. While some groups showed consistently poorer outcomes across all outcomes examined (e.g. low education, black non-Hispanic, unemployed/retired), the effects of some other predictors varied depending on the outcome. While these results bring to mind access to care issues (among Hispanics) as well as possibly lower availability and access to assistive devices (among women), they also suggest that poor health does not have to result in disability or poor quality of life. Future studies should examine the causes of such variation across outcomes, and develop potential workplace intervention strategies for improving health status of currently disadvantaged groups and enabling them to remain in the workforce.
Chapter IV: Specific Aim / Paper #2

I. Background

The number of adults aged 65 and over is not just increasing rapidly in the US population, but also constitute a quickly growing portion of the US workforce. Employment beyond age 65 is associated with better health and survival outcomes in older workers, while early retirement is associated with poorer health outcomes and greater mortality. While a strong healthy worker effect is apparent among older adults with healthier individuals continuing to work because their health allows them to, there are also potential intrinsic health benefits associated with employment at older age. For example, older workers’ health could be improved via a combination of increased physical activity associated with work, increased income, increased social engagement, and access to better health insurance coverage through employer, and these health benefits differ across older worker groups.

Health-related quality of life (HRQL) measures evaluate health status both in terms of the physical impact of ill health, as well as in terms of its wider context, such as emotional or social functioning, as perceived by the individual. HRQL is predictive of short-term and long-term mortality and health outcomes in both the general and older population. HRQL measures can be useful for identification of health disparities, and HRQL improvement is among the overarching goals of the Healthy People 2020. However, no systematic comparison of HRQL in subgroups of workers and non-workers aged 65+ has been undertaken in the US. In the current study, we examined the effect of work overall as well as work in four major occupational groups (white collar, blue collar, farming, and service) on HRQL of older adults using a representative sample of older US
adults. We hypothesized that workers would still have better HRQL scores, even after eliminating the effect of such health status predictors as the number of chronic health conditions, smoking status, and health insurance coverage.

II. Methods

i. Sample

The Medical Expenditure Panel Survey (MEPS) data were pooled for years 2000-2009 for adults aged 65+. MEPS is a nationally representative survey of the US non-institutionalized residents and their medical providers that gathers information on a variety of topics relating to health care costs, health service use, insurance coverage, access to care, and socio-demographic characteristics for all members of the participating households. Each household is followed for a period of 2 years with a total of 5 interviews (i.e. “rounds”) during this period. The Household component of MEPS is supplemented and corrected with the information collected through the Provider Component.

ii. Outcome Variables

Generic HRQL measures include health profile measures, which assess various dimensions of health individually, and utility measures, which combine information about different aspects of health into a single score. Examples of both groups are available through the Self-Administered Questionnaire (SAQ) of the MEPS Household component. These are the Short Form-12 (SF-12, available in years 2000-2009) and the EuroQol-5D (EQ-5D, available in years 2000-2003) measures,
respectively. The SAQ is a mail-back survey administered to all adults aged 18+ in the participating households once during each survey year (rounds 2 and 4).

**SF-12 v2.** In 2002, the SF-12 measure was updated and replaced with SF-12 version 2 [SF-12v2], which continued to be assessed through the following MEPS years. In this study, SF-12 scores were recoded into SF-12v2 scores as described in the SF-12 manual. This measure is based on 12 questions which separately assess the mental and the physical dimensions of the health-related quality of life. The scores for each of these dimensions are calculated separately into two subscales: the mental health component [MCS] and the physical health component of SF12 [PCS]. Scores for both subscales are designed to range from 0 to 100 with higher scores indicating better health, and these are scaled such that the mean of the general population is 50 with a standard deviation of 10.

**EQ-5D.** EQ-5D was assessed in MEPS during years 2000 through 2003. It is a 5-question measure that assesses 5 dimensions of functioning: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each of these dimensions is assessed with 3 different levels of response: no problem, mild problem, or severe problem. Based on these 5 questions, a preference score characterizing one’s health state is calculated, with higher scores indicating health states for which individuals have a stronger preference. The score ranges from 0 (health state equivalent in preference to death) to 1 (health state equivalent to perfect health). Negative EQ-5D scores are sometimes possible, indicating health states worse than death. Both self-reported scores as well as those reported by proxies were included in this study.
iii. Predictor variables

The effect of work on HRQL was assessed using a 5-category employment/occupation variable coded as: 1) unemployed/retired [reference], 2) white-collar workers, 3) blue-collar workers, 4) farm workers, and 5) service workers. Employment was assessed for the week prior to the same interview during which HRQL status was assessed (i.e., round 2 or 4). There was no distinction made between unemployed older individuals who have previously retired and those who did not work for some other reason. Both part time and full time workers were classified as “employed,” and occupation was assessed with a question about the kind of work the person did at the time of the interview.

The analyses were controlled for gender (male, female [reference]), race/ethnicity (non-Hispanic white [reference], Hispanic, non-Hispanic black, other), age in years, number of years of education, current smoking status (smoker [reference], non-smoker), health insurance status (private, public [reference], none), family income, and the number of co-existing chronic health conditions. Non-smoker group included former smokers, as it was not possible to know if the person has ever smoked previously. Family income was classified with respect to the national poverty level and was treated as an ordinal variable: poor (<100% of the poverty level), near poor (100%-125%), low income (125%-200%), middle income (200%-400%), and high income (>400% of the poverty level). The number of chronic conditions variable was a sum of self-reported life-time diagnoses of the following: asthma, hypertension, heart disease (included angina, myocardial infarction, coronary heart disease, or any other heart disease), diabetes, emphysema, stroke, and arthritis. The question about
arthritist diagnosis was only introduced into MEPS in 2001, and as this condition has a great potential effect on HRQL, chronic condition variable was only created for data years 2001-2009.\textsuperscript{126}

iv. Statistical Analysis

Figure 4.1 represents the overall conceptual model of the study. The associations between employment/occupation and HRQL variables were first modeled using multivariable regression adjusted for education, age, smoking status, gender, race/ethnicity chronic health conditions, income, and health insurance status in order to identify the variables that had a statistically significant association with the outcomes. Next, structural equation modeling (SEM) was used to estimate path models for each of the following outcomes: EQ-5D, MCS, and PCS. Two models
were estimated for each of the outcomes: one with a binary employment variable (employed/unemployed), and one with a 5-category employment/occupation variable, as described above. Education, age, smoking status, gender, and race/ethnicity were included as exogenous variables; chronic health conditions, income, health insurance status, and occupation were included as potential mediators. Statistically insignificant paths were removed from the models. Analyses were performed using a maximum likelihood estimator (MLR), with standard errors robust to non-independence and non-normality. MLR was chosen as it uses a logit function to model categorical mediators, and because it allows fullest use of available data such that observations with missing information are included in the analysis. Results were adjusted for sampling weights and design effects. The comparative fit of nested models was assessed using the likelihood ratio test. For the purposes of mediation assessment, ordinal and binary variables were treated as continuous. This is necessary for using a product of coefficients method of estimating indirect (mediated) pathways.

Descriptive analyses were performed using SAS 9.3 (SAS Institute Inc., Cary, NC). Model-based analyses were performed using Mplus version 6.1 (Muthen & Muthen 1998/2010).

III. Results

The pooled sample from 2000-2009, which was used for modeling SF-12 measures, included 34,643 participants representing approximately 37 million older US adults (Table 4.1). Of these, 13,816 participants from years 2000-2003 were included in the analysis modeling EQ-5D. The typical study participant was a 74 year old white non-

<table>
<thead>
<tr>
<th></th>
<th>Sample N</th>
<th>Estimated Population</th>
<th>Percent</th>
<th>Standard Error of %</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total Sex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>14,404</td>
<td>15,934,915</td>
<td>43.02</td>
<td>0.30</td>
<td>42.44 - 43.60</td>
</tr>
<tr>
<td>Female</td>
<td>20,239</td>
<td>21,104,100</td>
<td>56.98</td>
<td>0.30</td>
<td>56.40 - 57.56</td>
</tr>
<tr>
<td><strong>Income level</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poor</td>
<td>5,537</td>
<td>3,606,403</td>
<td>9.74</td>
<td>0.25</td>
<td>9.25 - 10.22</td>
</tr>
<tr>
<td>Near poor</td>
<td>2,676</td>
<td>2,462,026</td>
<td>6.65</td>
<td>0.20</td>
<td>6.25 - 7.05</td>
</tr>
<tr>
<td>Low income</td>
<td>6,551</td>
<td>7,045,874</td>
<td>19.02</td>
<td>0.38</td>
<td>18.29 - 19.76</td>
</tr>
<tr>
<td>Middle income</td>
<td>9,923</td>
<td>11,374,482</td>
<td>30.71</td>
<td>0.42</td>
<td>29.89 - 31.53</td>
</tr>
<tr>
<td>High income</td>
<td>9,956</td>
<td>12,550,229</td>
<td>33.88</td>
<td>0.61</td>
<td>32.68 - 35.08</td>
</tr>
<tr>
<td><strong>Insurance status</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Private</td>
<td>17,715</td>
<td>21,206,674</td>
<td>57.26</td>
<td>0.62</td>
<td>56.03 - 58.48</td>
</tr>
<tr>
<td>Public</td>
<td>16,700</td>
<td>15,683,124</td>
<td>42.34</td>
<td>0.62</td>
<td>41.12 - 43.56</td>
</tr>
<tr>
<td>None</td>
<td>228</td>
<td>149,217</td>
<td>0.40</td>
<td>0.05</td>
<td>0.31 - 0.49</td>
</tr>
<tr>
<td><strong>Previously retired</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>18,961</td>
<td>21,511,414</td>
<td>67.10</td>
<td>0.53</td>
<td>66.06 - 68.13</td>
</tr>
<tr>
<td>No</td>
<td>10,111</td>
<td>10,549,277</td>
<td>32.90</td>
<td>0.53</td>
<td>31.87 - 33.94</td>
</tr>
<tr>
<td><strong>Employment/occupation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unemployed/Retired</td>
<td>28,931</td>
<td>30,331,269</td>
<td>83.43</td>
<td>0.37</td>
<td>82.70 - 84.15</td>
</tr>
<tr>
<td>White Collar</td>
<td>3,069</td>
<td>3,880,536</td>
<td>10.67</td>
<td>0.33</td>
<td>10.02 - 11.32</td>
</tr>
<tr>
<td>Blue Collar</td>
<td>921</td>
<td>1,043,248</td>
<td>2.87</td>
<td>0.14</td>
<td>2.60 - 3.14</td>
</tr>
<tr>
<td>Farmer</td>
<td>108</td>
<td>94,339</td>
<td>0.26</td>
<td>0.04</td>
<td>0.17 - 0.35</td>
</tr>
<tr>
<td>Service</td>
<td>1,002</td>
<td>1,007,488</td>
<td>2.77</td>
<td>0.15</td>
<td>2.47 - 3.07</td>
</tr>
<tr>
<td><strong>Race/ethnicity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Hispanic white</td>
<td>23,951</td>
<td>30,006,427</td>
<td>81.01</td>
<td>0.61</td>
<td>79.81 - 82.22</td>
</tr>
<tr>
<td>Non-Hispanic black</td>
<td>4,767</td>
<td>3,083,365</td>
<td>8.33</td>
<td>0.38</td>
<td>7.58 - 9.07</td>
</tr>
<tr>
<td>Hispanic</td>
<td>4,239</td>
<td>2,381,823</td>
<td>6.43</td>
<td>0.40</td>
<td>5.64 - 7.22</td>
</tr>
<tr>
<td>Other</td>
<td>1,686</td>
<td>1,567,130</td>
<td>4.23</td>
<td>0.30</td>
<td>3.64 - 4.82</td>
</tr>
<tr>
<td><strong>Current smoker</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>3,716</td>
<td>3,796,909</td>
<td>10.54</td>
<td>0.29</td>
<td>9.97 - 11.11</td>
</tr>
<tr>
<td>No</td>
<td>29,887</td>
<td>32,229,680</td>
<td>89.46</td>
<td>0.29</td>
<td>88.89 - 90.03</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable means</th>
<th>Mean</th>
<th>SE of Mean</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of chronic health conditions</td>
<td>1.94</td>
<td>0.01</td>
<td>0 - 7</td>
</tr>
<tr>
<td>Number of years of education</td>
<td>12.10</td>
<td>0.04</td>
<td>0 - 17</td>
</tr>
<tr>
<td>Age</td>
<td>74.40</td>
<td>0.07</td>
<td>65 - 85</td>
</tr>
<tr>
<td>SF-12 physical component score</td>
<td>41.73</td>
<td>0.11</td>
<td>5.85 - 69.22</td>
</tr>
<tr>
<td>SF-12 mental component score</td>
<td>51.91</td>
<td>0.09</td>
<td>0.76 - 77.37</td>
</tr>
<tr>
<td>EQ-5D score</td>
<td>0.72</td>
<td>0.00</td>
<td>-0.59 - 1.00</td>
</tr>
</tbody>
</table>

*a* Information about previous retirement was not available for a large proportion of unemployed participants
Hispanic (81%) female (57%) non-smoker (89%) with a high school education or equivalent (mean of 12 years of education). The majority came from households with middle or high income levels (65%), and less than a half a percent had no health insurance. The majority did not work at the time of the interview (83%); however only 67% of the participants reported having retired. On average, participants had a PCS score of 42 (range: 5.85 - 69.22), MCS of 52 (range: 0.76 - 77.37), and EQ-5D of 0.72 (range: -0.59 - 1.00).

Table 4.2 presents unstandardized coefficients for the multivariable regression and path models. None of the predictors became insignificant in path models when compared to regression models. The largest significant effect on all HRQL outcomes was produced by the number of chronic health conditions: in the path models, for each additional condition, MCS scores decreased by -2.20 points [95% Confidence Interval: -2.52; -1.88], PCS decreased by -6.63 [-6.99; -6.28], and EQ5D decreased by -0.10; [-0.11; -0.08]. Employment had the next greatest magnitude of association with PCS (3.71; [4.22, 3.19]) and EQ5D (0.04; [0.06, 0.03]), while current smoking (-1.91; [-2.42; -1.41]) and Hispanic and other race/ethnicity were the next strongest predictors of MCS (-1.79; [-2.38; -1.21] and -1.59; [-2.39; -0.78], respectively). Race/ethnicity had no effect on EQ5D or PCS scores after controlling for the effect of all other covariates. Among workers, farm workers had higher MCS scores than white collar workers in multivariable regression analyses (2.32; [0.07; 4.57]). However this effect was no longer significant in the path model, suggesting that income and health insurance coverage partially mediated this relationship. In path models, blue collar workers had slightly better MCS scores than white collar workers (1.08; [0.22; 1.93]). Both farm and service workers had higher
Table 4.2: Health Related Quality of Life Regression Results and Select Path Model Results, Medical Expenditure Panel Survey 2000-2009 Participants Aged 65+.

<table>
<thead>
<tr>
<th></th>
<th>Model 1: Multivariable Regression Analysis</th>
<th>Model 2: Full Path Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MCS</td>
<td>PCS</td>
</tr>
<tr>
<td></td>
<td>Estimate</td>
<td>95% CI</td>
</tr>
<tr>
<td>Work status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employed vs. unemployed/retired</td>
<td>1.49</td>
<td>1.02, 1.97</td>
</tr>
<tr>
<td>Occupation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unemployed vs. white collar</td>
<td>-1.217</td>
<td>-1.78, -0.65</td>
</tr>
<tr>
<td>Blue collar vs. white collar</td>
<td>0.92</td>
<td>-0.03, 1.87</td>
</tr>
<tr>
<td>Farm worker vs. white collar</td>
<td>2.32</td>
<td>0.07, 4.57</td>
</tr>
<tr>
<td>Service worker vs. white collar</td>
<td>0.43</td>
<td>-0.58, 1.43</td>
</tr>
<tr>
<td>Covariate controls</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female vs. male</td>
<td>-0.72</td>
<td>-1.05, -0.39</td>
</tr>
<tr>
<td>Income level</td>
<td>0.51</td>
<td>0.38, 0.65</td>
</tr>
<tr>
<td>Insurance: private vs. public</td>
<td>1.15</td>
<td>0.73, 1.57</td>
</tr>
<tr>
<td>Chronic conditions</td>
<td>-2.19</td>
<td>-2.50, -1.87</td>
</tr>
<tr>
<td>Age</td>
<td>-0.03</td>
<td>-0.06, 0.00</td>
</tr>
<tr>
<td>Education years</td>
<td>0.27</td>
<td>0.21, 0.33</td>
</tr>
<tr>
<td>Smoker vs. non-smoker</td>
<td>-1.63</td>
<td>-2.18, -1.07</td>
</tr>
<tr>
<td>Black vs. white non-Hispanic</td>
<td>-0.11</td>
<td>-0.75, 0.52</td>
</tr>
<tr>
<td>Hispanic vs. white non-Hispanic</td>
<td>-1.79</td>
<td>-2.48, -1.02</td>
</tr>
</tbody>
</table>

Notes: EQ-5D analysis only included data for years 2000-2003. MCS: mental health component summary of SF-12. PCS: physical health component summary of SF-12. Number of chronic conditions was transformed using quadratic root.
EQ5D scores than white collar workers (0.05; [0.01; 0.09] and 0.03; [0.01; 0.06], respectively). Compared to white collar workers, all other workers had similar PCS scores. In addition, female gender, increased age, and current smoking were each associated with lower quality of life as measured by each of the three HRQL measures. Higher income levels, private health insurance coverage, and higher education were associated with higher HRQL scores across all three outcome measures.

The standardized estimates for select mediation effects are presented in Table 4.3, and direct standardized path estimates are presented in Figure 4.2. The indirect effects were small in magnitude, however highly statistically significant. All the indirect effects
<table>
<thead>
<tr>
<th></th>
<th>MCS estimate</th>
<th></th>
<th>PCS estimate</th>
<th>p-value</th>
<th>EQ-5D estimate</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Current employment</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direct effects</td>
<td>0.04</td>
<td>&lt;0.001</td>
<td>0.10</td>
<td>&lt;0.001</td>
<td>0.04</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Indirect via income</td>
<td>0.01</td>
<td>&lt;0.001</td>
<td>0.01</td>
<td>&lt;0.001</td>
<td>0.01</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Total effect</td>
<td>0.06</td>
<td>&lt;0.001</td>
<td>0.11</td>
<td>&lt;0.001</td>
<td>0.05</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Chronic health conditions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direct effects</td>
<td>-0.16</td>
<td>&lt;0.001</td>
<td>-0.41</td>
<td>&lt;0.001</td>
<td>-0.33</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Indirect via employment</td>
<td>0.00</td>
<td>&lt;0.001</td>
<td>-0.01</td>
<td>&lt;0.001</td>
<td>0.00</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Total effect</td>
<td>-0.17</td>
<td>&lt;0.001</td>
<td>-0.42</td>
<td>&lt;0.001</td>
<td>-0.33</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Income level</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direct effects</td>
<td>0.07</td>
<td>&lt;0.001</td>
<td>0.04</td>
<td>&lt;0.001</td>
<td>0.06</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Indirect via insurance coverage</td>
<td>0.01</td>
<td>&lt;0.001</td>
<td>0.01</td>
<td>&lt;0.001</td>
<td>0.01</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Total effect</td>
<td>0.08</td>
<td>&lt;0.001</td>
<td>0.05</td>
<td>&lt;0.001</td>
<td>0.07</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Female gender</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direct</td>
<td>-0.02</td>
<td>0.001</td>
<td>-0.04</td>
<td>&lt;0.001</td>
<td>-0.05</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Indirect via income</td>
<td>-0.01</td>
<td>&lt;0.001</td>
<td>0.00</td>
<td>&lt;0.001</td>
<td>-0.01</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Indirect via chronic</td>
<td>0.00</td>
<td>0.016</td>
<td>-0.01</td>
<td>0.013</td>
<td>-0.01</td>
<td>0.216</td>
</tr>
<tr>
<td>Indirect via employment</td>
<td>0.00</td>
<td>&lt;0.001</td>
<td>-0.01</td>
<td>&lt;0.001</td>
<td>0.00</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Total</td>
<td>-0.04</td>
<td>&lt;0.001</td>
<td>-0.06</td>
<td>&lt;0.001</td>
<td>-0.06</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Current smoking</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direct</td>
<td>-0.06</td>
<td>&lt;0.001</td>
<td>-0.03</td>
<td>0.002</td>
<td>-0.04</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Indirect via chronic</td>
<td>0.00</td>
<td>0.009</td>
<td>0.01</td>
<td>0.013</td>
<td>0.01</td>
<td>0.010</td>
</tr>
<tr>
<td>Total</td>
<td>-0.06</td>
<td>&lt;0.001</td>
<td>-0.04</td>
<td>&lt;0.001</td>
<td>-0.03</td>
<td>0.004</td>
</tr>
</tbody>
</table>
listed in this section were of magnitude 0.01 SD units and significant at p<0.001 level unless noted otherwise. Employment via its effect on income was significantly associated with higher EQ5D, MCS, and PCS scores, however this effect was a much smaller proportion of the total effect for PCS (total effect: 0.11 SD, p<0.001) than for the other two outcomes. Among the employed, there were no significant differences in mediation effects between occupations. Higher income was associated with higher EQ5D, MCS, and PCS scores via its effect on health insurance coverage. The effect of the number of chronic health conditions was mostly direct, with only a very small indirect negative effect on PCS via decreased employment (-0.01 SD). Female gender was associated with decreased MCS and EQ5D scores via income, decreased scores on PCS and EQ5D via the number of chronic conditions, and with decreased PCS scores via its effect on employment (all -0.01 SD). Current smoking was associated with increased PCS and EQ5D scores via its effect on the number of chronic health conditions. All other indirect effects had only a minor contribution to total effects (<0.01 SD units).

IV. Discussion

In this study, we examined the association of employment and occupation, as well as other socio-demographic and health predictors, with three different HRQL outcomes of older adults: EQ-5D and the mental and physical components of SF-12 (i.e. MCS and PCS). By comparison to MCS and PCS (measures that each assess a separate dimension of HRQL - mental and physical), EQ-5D is a measure that includes both mental and physical functioning.
Overall, the findings were consistent across most socio-demographic measures for all three outcomes with the exception of race/ethnicity, which was strongly associated with MCS scores only, potentially indicative of differential item functioning between racial/ethnic groups.\textsuperscript{169} The number of chronic health conditions was associated with the greatest change in all HRQL outcome measures (range: 0.16-0.41 SD decrease), with almost all of the effect being direct. The conditions included in this count variable varied from those of potentially low impact on HRQL (e.g. asthma) to those associated with substantial worsening in HRQL scores (e.g. diabetes, arthritis).\textsuperscript{126,127} These were treated equally in estimating the model, however different chronic health conditions might not have the same detrimental effect on HRQL. Of all outcome measures, chronic conditions explained the most variance in PCS and the smallest variance in MCS. This is consistent with previous MEPS findings of a much stronger effect size of chronic conditions on PCS than on MCS.\textsuperscript{170}

After the chronic condition count, employment was the next strongest predictor of PCS and EQ5D, and it was strongly associated with improved MCS scores as well. However, Hispanic race/ethnicity and current smoking were associated with a greater negative change in MCS scores. Employment also had the strongest association with PCS, which highlights the importance of good physical health for employment at older age. There was no difference between occupations in PCS scores suggesting that older workers of all occupations may be in similarly good physical health, enabling them to continue working. There was however a variation in the MCS and EQ5D scores across occupations. Farm work was the strongest predictor of higher MCS scores in multivariable regression analysis, however that effect disappeared after introducing
mediation pathways, possibly due to reduced power and small sample sizes in this occupational group. In addition, blue collar work became a significant predictor of high MCS scores in a path model when compared to white collar workers. For EQ5D, farm and service workers had higher scores than white collar workers.

Income partially mediated some of the effect of employment on HRQL, however this effect was weaker for PCS scores relative to the direct effect of employment. While employment was associated with an even greater improvement in PCS scores than in other measures, most of it was due to the direct effect of employment. This again is consistent with the healthy worker effect, but also could result from increased physical activity associated with employment. Income itself was associated with a more modest increase in PCS scores than in other HRQL measures, with some of its effect on all HRQL outcomes partially mediated by more comprehensive health insurance coverage.

The negative effect of female gender on MCS and EQ5D was partially mediated by decreased income, however income did not contribute to the effect of gender on PCS. Instead, the negative effect of female gender on PCS was partially mediated by the increased number of chronic conditions and decreased employment rates among females. While women are less likely to be employed, they are also more likely to earn less than men,\textsuperscript{171-173} as well as to report multiple chronic health conditions.\textsuperscript{174} While current smoking also had an overall negative effect on all HRQL measures, some of that effect on PCS and EQ5D was attenuated by the lower number of chronic health conditions among current smokers.

While the EQ5D outcomes mirrored those for SF-12, there were some inconsistencies, notably in modeling the effects of race/ethnicity and occupation.
Previously, SF-12 was shown to be more sensitive at the highest HRQL values than EQ-5D\textsuperscript{161}. That is, even when individuals report perfect health on EQ-5D measures, slight variations in health can still be detected using SF-12. This might explain some of the differences in results, however not the significantly higher EQ5D scores among farm and service workers. While models mapping EQ5D and SF-12 measures onto each other have been developed,\textsuperscript{175,176} these do not map onto each other perfectly.\textsuperscript{177} In this study, the use of both measures allowed to capture slightly different aspects of the HRQL.

A previous nationally-representative study using National Health Interview Survey data linked to MEPS has found the EQ5D scores to be the highest in white collar workers,\textsuperscript{178} while we found farmers and service workers to have higher scores than white collar. In the above study, EQ5D scores were measured 1-2 years after the employment/occupation status was assessed, and therefore for the older participants the occupation information might not have been accurate. However, these differences might also indicate a more rapid deterioration in EQ5D scores among the farmers and service workers leading to lower scores at 1-2 years follow-up even with higher baseline scores. Future studies should examine these effects using longitudinal data.

V. Limitations

Because MLR estimator was used for modeling the data, no standard model fit indices (such as RMSEA, CFI, etc.) could be obtained. We attempted to obtain model fit indices on a subsample of the data using the weighted least squares (WLS) estimator. However, WLS uses a probit function to estimate categorical mediators, and it uses listwise deletion for the treatment of missing data; therefore, the resulting estimates were
very different from MLR estimates. Categorical mediators also had to be treated as continuous for the purposes of indirect effect estimation. This results in approximate effect size estimation where accuracy is diminished to the degree that the categorical variable deviates from a normal distribution. In addition, the cross-sectional data used in this study did not allow to make causal inferences, and therefore to really accurately assess mediation.

VI. Conclusions

While adjustment for co-morbidities and smoking status in our study has potentially eliminated some of the healthy-worker effect, employment was still the strongest predictor of higher HRQL scores, especially on those measures that contained a mental health component, and especially in more physically demanding occupations. Some of this effect on mental health was mediated by income. Social engagement associated with work is one explanation for improved mental health scores, and it might be more valuable for lower qualified individuals engaged in physical work. For these individuals some of the mental health benefits might also be associated with being able to earn a living. There was no difference between workers in physical health measures, however employment was strongly associated with physical health, and most of the effect was direct. Physical activity associated with work might lead to improved physical health even at relatively low activity levels (e.g. in office workers).

Individuals who are not healthy enough for blue collar jobs but are not qualified enough for white collar positions might therefore lack access to the benefits associated with employment. Better workplace accommodations for older adults with limitations, as
well as workplace chronic disease management programs, could allow older adults in poorer physical health access to a greater variety of jobs and to benefits of employment. These benefits could be associated both with employment itself, as well as with increased income and better health insurance coverage, as these factors partially mediated the effects of employment in our study. Further understanding of the effect of work on health at older age could assist in the development of policy measures aimed at improving older adults’ HRQL.
Chapter V: Specific Aim / Paper #3

I. Background

The older population of the developed world has been expanding dramatically over the recent years.\textsuperscript{2,114,179} In the US and other developed nations, adults aged over 65 account for the largest proportion of healthcare costs.\textsuperscript{73,74} To reduce healthcare costs, it is important to identify modifiable factors among this population.

For the current older population, early or life-time risk factor exposures, such as low education, history of occupational exposures, and smoking, are already too late to address. Nevertheless, identifying those populations that are most at risk, as well as those populations least at risk, will allow for better targeting of health promotion measures. Among potential modifiable factors affecting costs is employment among older workers. While healthy people are more likely to continue employment past retirement age, working beyond age 65 can also contribute to overall physical activity and social engagement of older adults, and as a result can lead to improved health and lower healthcare costs.\textsuperscript{18,19,162}

We explore healthcare expenditures among older adults, as well as the use of two healthcare services that are likely to lead to unnecessary and preventable increases in cost: the number of visits to the emergency room (ER), and the number of nights spent in a hospital.\textsuperscript{180,181} We assess the effects of both the predictors that are hypothesized to increase the costs and utilization (e.g. the number of chronic conditions, smoking) as well as those that are hypothesized to decrease expenditures and utilization (e.g. higher education and income, employment) on these outcomes using nationally representative
data and controlling for the effects of socio-demographic characteristics (e.g. gender, age, race/ethnicity).

II. Methods

i. Sample

Data from the Medical Expenditure Panel Survey (MEPS) were pooled for years 2000-2009 for adults aged 65+. The MEPS is a yearly panel survey designed to be representative of the non-institutionalized US population. In addition to interviewing household members, MEPS also collects data from their medical providers and employers. Detailed information is collected through a series of 5 interviews (i.e. “rounds”) conducted over 2 years on each member of the participating households’ socio-demographic characteristics, health service use and associated costs, access to care, and health insurance coverage. The information collected from the households is corrected and supplemented using the information obtained through health service providers to yield the Household Component of MEPS, utilized in this study. This study was approved by the University of Miami Institutional Review Board. The first author takes complete responsibility for the integrity of the data and the accuracy of the data analysis.

ii. Outcome Variables

Detailed data on health care expenditures and health service utilization are collected through MEPS. These are summed over each survey year for each
participant, and the cumulative values are available in the full-year consolidated MEPS files.

**Healthcare Utilization.** Utilization event information is collected through the Household component of the MEPS. Utilization events are summed for each individual within several categories: medical provider visits, both office-based and hospital outpatient; hospital inpatient stays and the number of nights per stay; ER visits; dental visits; prescription medicine purchases; home health care; and other medical equipment and services utilization events. We limited the scope of this study to the events that are likely to result in high and preventable costs: ER visits (quantified as the annual number of visits) and hospital admissions (quantified as the annual number of nights of hospital stay).

**Health Care Expenditures.** Expenditure information in MEPS is collected using both the household interviews and the medical provider surveys, with the provider information serving as the primary source of expenditure data for most categories of expenditures. This information represents the sum of payments made to providers from all sources (i.e. out-of-pocket, insurance). Total expenditure amounts for the calendar year were used as an outcome in this study. These were converted into 2010 dollars using the Bureau of Labor Statistics consumer price index values.20

### iii. Predictor variables

The following variables were included as predictors in all models:

- employment/occupation (unemployed/retired, white-collar workers [reference group], blue-collar workers, farm workers, and service workers),
- gender (male [reference],
female), age at the beginning of the year, years of education, race/ethnicity (non-Hispanic white [reference], Hispanic, non-Hispanic black, other), family income (see description below), smoking status (current smoker, current non-smoker [reference]), health insurance status (private, public [reference], none), and the number of reported chronic health conditions.

Occupation and employment were assessed at the beginning of the year for which expenditures and utilization were measured (i.e. MEPS round 1 or 3) with questions about whether the person worked in the previous week, and what kind of work they did. Previous retirement information was not assessed for a large proportion of the participants, and therefore no distinction was made in this study between retired individuals and those who did not work for some other reason. It was also not possible to know what the individual’s earlier occupations might have been. The employed categories included both part-time and full-time workers.

Family income was categorized as a percentage of the national poverty level into: 1) poor (<100%), near poor (100%-125%), low income (125%-200%), middle income (200%-400%), and high income (>400%). The count of chronic conditions was a sum of life-time diagnoses of the following: asthma, hypertension, heart disease (included angina, myocardial infarction, coronary heart disease, or any other heart disease), diabetes, emphysema, stroke, and arthritis. Health insurance status was classified according to the greatest level of coverage the person had within that year. Public insurance included Medicare, Medicaid, and other public hospital/physician coverage. Private insurance plans (including Medigap and TRICARE) were classified as such if they provided at a minimum hospital and physician service coverage. Due
to this population’s eligibility for Medicare, sensitivity analyses were performed using only the insured individuals and excluding the 0.4% that were uninsured.

Figure 5.1. Full Path Model Predicting Healthcare Expenditures and Utilization

iv. Statistical Analysis

The overall conceptual model of the study is presented in Figure 5.1. We used structural equation modeling to estimate path models for each of the outcomes. Poisson regression was used for modeling ER visits. Negative binomial regression was used for modeling expenditures and the nights of hospital stay due to greater dispersion in these outcomes. Due to a large number of zero values in all three
outcomes, all models were zero-inflated. In the path models, the following were modeled as exogenous variables: gender, race/ethnicity, education, age, and smoking status. The following variables were modeled as mediators: number of chronic health conditions, income, health insurance status, and employment/occupation. Paths that were not statistically significant at p=0.05 level were removed from the models.

Analyses utilized the maximum likelihood estimator (MLR) with robust standard errors. MLR uses a logit function to model categorical mediators and allows fullest use of available data by retaining observations with missing data. Mediation was tested using a product of coefficients method, however only the direct effects are discussed in this study. Analyses were adjusted for sampling weights and complex survey design effects. SAS 9.3 (SAS Institute Inc., Cary, NC) was used for all descriptive analyses, and Mplus version 6.1 (Muthen & Muthen 1998/2010) was used for the structural equation model-based analyses.

III. Results

i. Sample Characteristics

A total of 34,276 yearly observations from 19,720 participants aged 65+ were included in the study, which represented approximately 37 million of older US adults annually (Table 5.1). The results obtained using the sample of insured older adults only were not different from results of the complete sample, therefore the results from the complete sample analyses are presented. The majority of the participants were unemployed/retired (82%), with the employed participants being on average younger (mean age range: 69.4 – 70.49) and having fewer chronic health conditions (mean
<table>
<thead>
<tr>
<th></th>
<th>Unemployed/retired</th>
<th>White Collar</th>
<th>Blue Collar</th>
<th>Farm</th>
<th>Service</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MEPS sample</td>
<td>Estimated</td>
<td>Percent</td>
<td>MEPS sample</td>
<td>Estimated</td>
</tr>
<tr>
<td>Total</td>
<td>28,921</td>
<td>30,397,622</td>
<td>81.99</td>
<td>959</td>
<td>1,079,661</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>11,312</td>
<td>12,275,663</td>
<td>40.42</td>
<td>818</td>
<td>929,313</td>
</tr>
<tr>
<td>Female</td>
<td>17,609</td>
<td>18,091,961</td>
<td>59.58</td>
<td>112</td>
<td>150,347</td>
</tr>
<tr>
<td>Income level</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poor</td>
<td>5,339</td>
<td>3,478,793</td>
<td>11.46</td>
<td>24</td>
<td>16,235</td>
</tr>
<tr>
<td>Near poor</td>
<td>2,558</td>
<td>2,384,976</td>
<td>7.76</td>
<td>19</td>
<td>15,492</td>
</tr>
<tr>
<td>Low income</td>
<td>5,933</td>
<td>6,420,098</td>
<td>21.14</td>
<td>207</td>
<td>257,843</td>
</tr>
<tr>
<td>Middle income</td>
<td>8,205</td>
<td>9,443,908</td>
<td>31.10</td>
<td>799</td>
<td>957,571</td>
</tr>
<tr>
<td>High income</td>
<td>6,896</td>
<td>8,628,855</td>
<td>28.55</td>
<td>2,109</td>
<td>2,741,794</td>
</tr>
<tr>
<td>Insurance coverage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Private</td>
<td>13,868</td>
<td>16,575,510</td>
<td>54.52</td>
<td>640</td>
<td>756,858</td>
</tr>
<tr>
<td>Public</td>
<td>14,906</td>
<td>13,992,315</td>
<td>45.09</td>
<td>842</td>
<td>1,059,869</td>
</tr>
<tr>
<td>None</td>
<td>177</td>
<td>117,800</td>
<td>0.39</td>
<td>15</td>
<td>10,939</td>
</tr>
<tr>
<td>Race / ethnicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Hispanic white</td>
<td>19,866</td>
<td>24,583,651</td>
<td>80.89</td>
<td>2,597</td>
<td>3,502,008</td>
</tr>
<tr>
<td>Non-Hispanic black</td>
<td>4,013</td>
<td>2,528,006</td>
<td>8.32</td>
<td>273</td>
<td>215,588</td>
</tr>
<tr>
<td>Hispanic</td>
<td>3,662</td>
<td>1,972,947</td>
<td>6.50</td>
<td>186</td>
<td>108,282</td>
</tr>
<tr>
<td>Other</td>
<td>1,410</td>
<td>1,303,021</td>
<td>4.29</td>
<td>153</td>
<td>152,087</td>
</tr>
<tr>
<td>Current smoker</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>24,976</td>
<td>26,453,464</td>
<td>87.11</td>
<td>2,855</td>
<td>3,593,657</td>
</tr>
<tr>
<td>Yes</td>
<td>3,077</td>
<td>3,046,539</td>
<td>10.03</td>
<td>285</td>
<td>367,210</td>
</tr>
<tr>
<td>Chronic conditions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>0 - 7</td>
<td>2.02 (0.01)</td>
<td>0 - 6</td>
<td>1.54 (0.03)</td>
<td>0 - 6</td>
</tr>
<tr>
<td>Years of education</td>
<td>0 - 17</td>
<td>11.87 (0.05)</td>
<td>0 - 17</td>
<td>14.13 (0.06)</td>
<td>0 - 17</td>
</tr>
<tr>
<td>Age</td>
<td>65 - 85</td>
<td>75.31 (0.08)</td>
<td>65 - 85</td>
<td>70.25 (0.14)</td>
<td>65 - 85</td>
</tr>
<tr>
<td>Total expenditures</td>
<td>0 - 558,890</td>
<td>9,337 (124.38)</td>
<td>0 - 248,505</td>
<td>6,687 (246.05)</td>
<td>0 - 153,936</td>
</tr>
<tr>
<td>Emergency room visits</td>
<td>0 - 13</td>
<td>0.29 (0.01)</td>
<td>0 - 10</td>
<td>0.17 (0.01)</td>
<td>0 - 5</td>
</tr>
<tr>
<td>Nights for hospital stays</td>
<td>0 - 300</td>
<td>1.7 (0.05)</td>
<td>0 - 163</td>
<td>0.78 (0.09)</td>
<td>0 - 90</td>
</tr>
</tbody>
</table>
range 1.5-1.6) than their unemployed/retired counterparts (mean age: 75.3; 2.0 conditions on average).

### Table 5.2: Zero-Inflated Negative Binomial and Negative Poisson Regression Results, Direct Effects.

<table>
<thead>
<tr>
<th>Covariates</th>
<th>Model 1: Modeling Total Expenditures*</th>
<th>Model 2: Modeling Emergency Room Visits</th>
<th>Model 3: Modeling # of Nights for Hospital Stays</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR 95% CI</td>
<td>OR 95% CI</td>
<td>OR 95% CI</td>
</tr>
<tr>
<td>Female</td>
<td>1.47 (1.23; 1.76)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Income (5-category ordinal variable)</td>
<td>1.08 (1.01; 1.16)</td>
<td>0.90 (0.87; 0.93)</td>
<td>NA</td>
</tr>
<tr>
<td>Health insurance (Private vs Public)</td>
<td>1.84 (1.53; 2.23)</td>
<td>1.01 (0.91; 1.11)</td>
<td>1.04 (0.96; 1.13)</td>
</tr>
<tr>
<td>Health insurance (None vs Public)</td>
<td>0.19 (0.11; 0.34)</td>
<td>0.29 (0.12; 0.69)</td>
<td>0.40 (0.19; 0.87)</td>
</tr>
<tr>
<td>Number of chronic conditions</td>
<td>4.22 (3.72; 4.8)</td>
<td>1.28 (1.19; 1.37)</td>
<td>1.62 (1.55; 1.68)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>1.03 (1.02; 1.05)</td>
<td>1.03 (1.03; 1.04)</td>
<td>1.04 (1.03; 1.05)</td>
</tr>
<tr>
<td>Education (years)</td>
<td>1.13 (1.1; 1.16)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Black Non-Hispanic (vs White)</td>
<td>0.54 (0.4; 0.71)</td>
<td>NA</td>
<td>0.95 (0.84; 1.08)</td>
</tr>
<tr>
<td>Hispanic (vs White)</td>
<td>0.90 (0.68; 1.18)</td>
<td>NA</td>
<td>0.82 (0.7; 0.96)</td>
</tr>
<tr>
<td>Other race/ethnicity (vs White)</td>
<td>0.67 (0.43; 1.04)</td>
<td>NA</td>
<td>0.62 (0.49; 0.78)</td>
</tr>
<tr>
<td>Smoker (vs Non-smoker)</td>
<td>0.50 (0.4; 0.62)</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

|                                          | IRR  CI                                | IRR  CI                                | IRR  CI                                       |
|                                          |                                       |                                       |                                               |
| Health insurance (Private vs Public)     | 1.10 (1.05; 1.16)                      | NA                                    | NA                                            |
| Health insurance (None vs Public)        | 0.34 (0.24; 0.49)                      | NA                                    | NA                                            |
| Blue Collar (vs White Collar)            | 1.15 (1.07; 1.24)                      | 1.22 (1.06; 1.42)                     | 1.26 (1.01; 1.58)                             |
| Farmer (vs White Collar)                 | 1.05 (0.72; 1.53)                      | 1.48 (0.92; 2.39)                     | 1.32 (0.52; 3.38)                             |
| Service (vs White Collar)                | 0.77 (0.66; 0.89)                      | 1.05 (0.84; 1.31)                     | 1.14 (0.76; 1.7)                              |
| Number of chronic conditions             | 1.40 (1.38; 1.43)                      | 1.23 (1.17; 1.29)                     | 1.14 (1.09; 1.2)                              |
| Age (years)                              | 1.01 (1.01; 1.02)                      | NA                                    | NA                                            |
| Education (years)                        | 1.01 (1.01; 1.02)                      | NA                                    | 0.98 (0.96; 0.99)                             |
| Black Non-Hispanic (vs White)            | 0.98 (0.91; 1.07)                      | NA                                    | 1.24 (1.05; 1.47)                             |
| Hispanic (vs White)                      | 1.03 (0.94; 1.13)                      | NA                                    | 0.98 (0.81; 1.18)                             |
| Other race/ethnicity (vs White)          | 0.87 (0.77; 0.99)                      | NA                                    | 1.11 (0.75; 1.64)                             |
| Smoker (vs Non-smoker)                   | 0.91 (0.84; 0.98)                      | NA                                    | NA                                            |
| Female                                  | NA                                    | NA                                    | 0.86 (0.78; 0.96)                             |
| Income (5-category ordinal variable)     | NA                                    | NA                                    | 0.91 (0.88; 0.95)                             |

*Expenditures were calculated in 2010 dollars. OR: odds ratio; IRR: incidence rate ratio. NA – these pathways were removed from the models due to non-significance.
Over a half of those employed worked in white collar occupations (10.9% of total sample). Women comprised the majority of the unemployed/retired (59.6%) and of the service workers (58.8%), gender proportions were equal for the white collar workers, while the majority of blue collar and farm workers were male (86.0% and 80.6%, respectively). The unemployed/retired had the highest proportion of individuals fall into the poor or near poor category (19.2% vs 2.0% - 6.5% for workers). The greatest proportion of the uninsured was among the farm workers (2.26%), and the lowest was among the white collar workers (0.27%). White collar and blue collar occupations had the highest proportions of workers with private insurance coverage (73.5 and 70.1% respectively). The majority of the participants across all groups were non-Hispanic white, with the greatest proportion of older Hispanics among farm workers (17.5%), and the greatest proportion of non-Hispanic blacks among service workers (17.1%). While the mean total medical expenditures were similar across occupations (range: $5,300-$6,700), maximum expenditures within the occupations ranged from approximately $64,000 for farmers to $248,000 for white collar workers.

ii. Path Modeling: Probability of an Outcome

Table 5.2 presents the direct effect results of the zero-inflated Poisson and negative binomial modeling. Such modeling generates 2 sets of outcomes: one estimating the effect of the independent variables on the likelihood of the outcome being different from zero, e.g., having any expenditure (top part of the table, odds ratios); the other predicting the effect of the independent variables on the magnitude
of the outcomes adjusting for the probably of having an outcome (bottom part of the table, incidence rate ratios). Overall, older participants with more chronic conditions were more likely to have non-zero outcomes. Each additional chronic condition was associated with over 4 times the odds of having a health-related expenditure (Odds Ratio: 4.22; 95% Confidence Interval: 3.72-4.80), 62% greater odds of staying in a hospital overnight (1.62; 1.55-1.68), and 28% greater odds of ending up in an ER (1.28; 1.19; 1.37). Greater income was associated with lower odds of an ER visit (0.90; 0.87-0.93), however greater odds of having an expenditure (1.08; 1.01-1.16). Hispanics were less likely than non-Hispanic whites to spend time in a hospital (0.82; 0.70-0.96), however were no different from whites on the odds of having an expenditure or going to the ER. Being female (1.47; 1.23-1.76) and having private health insurance coverage (1.84; 1.53-2.23) were also associated with a greater odds of having an expenditure. Current smoking (0.50; 0.4-0.62) and black non-Hispanic race/ethnicity vis-a-vis whites (0.54; 0.4-0.71) were associated with lower odds of an expenditure. Employment and occupation were not associated with the odds of having an expenditure, ER visit, or a hospital stay.

iii. Path Modeling: Outcome Magnitude

For those participants who reported having an outcome (i.e. healthcare expenses, ER visit, or hospital stay), a greater number of ER visits was predicted by having an additional chronic health condition (Incidence Rate Ratio: 1.23; 95% Confidence Interval: 1.17-1.29). Individuals had higher expenditures if they had private health insurance coverage (1.10; 1.05-1.16), more chronic health conditions
(1.40; 1.38-1.43), were older (1.01; 1.01-1.02) and more educated (1.01; 1.01-1.02) had. Those with no health insurance (0.34; 0.24-0.49) and current smokers (0.91; 0.84-0.98) had lower expenditures. Greater number of nights in a hospital was spent by those with more chronic health conditions (1.14; 1.09-1.2) and of black non-Hispanic race/ethnicity (1.24; 1.05-1.47), while those more educated (0.98; 0.96-0.99), female (0.86; 0.78-0.96), and with greater income (0.91; 0.88-0.95) spent fewer nights in a hospital.

Older unemployed/retired individuals, when compared to white collar employees, experienced higher expenditures: (1.15; 1.07-1.24), more annual ER visits (1.22; 1.06-1.42), and spent more nights in a hospital (1.26; 1.01-1.58). Older service workers had lower expenditures as compared to white collar workers (0.77; 0.66-0.89). There were no differences between the occupations on other measures. There was a trend towards higher expenditures and healthcare utilization among farmers, which is noteworthy due to a small sample size in this group (n=105).

iv. Mean Expenditures

Table 5.3 presents the mean healthcare expenditures in 2010 dollars for each group of older adults, calculated with and without adjustment for the probability of having an expenditure. Mean expenditures for the average older adult (i.e. 74-year old non-smoker, with 2 chronic health conditions, 12 years of education, and Medicare or other public health insurance coverage) employed in a white collar occupation, after adjustment for the probability of expenditure, were $6,899. The highest expenditures were associated with having an additional chronic health condition ($9,664 for the
were among the uninsured ($2,364), with service workers ($5,314) and smokers ($6,249) being the next lowest cost groups. Other differences were either small in magnitude (e.g. age, additional years of education) or were not statistically significant in regression modeling. Differences between means in Table 5.3 were not tested for significance.

### IV. Discussion

In this study, we examined the factors affecting health-related expenditures and healthcare utilization (ER visits and number of nights spent in the hospital) among older...
US adults, both those employed and unemployed. We found that the unemployed/retired individuals were poorer, older, and in worse health, and they generally spent a lot more on healthcare than their peers who still worked. Among workers, farmers had the highest proportion of the uninsured and Hispanics, and also had the highest mean number of annual ER visits. White collar and blue collar workers were the most likely to have private insurance coverage. Service workers had the highest proportion of non-Hispanic blacks among them.

In agreement with previous studies, we found that the number of chronic conditions had a great effect on healthcare expenditures. We also found that after controlling for all other direct and indirect effects, the number of chronic conditions was the single most important factor in predicting healthcare costs and utilization; and that after controlling for the effect of chronic conditions, employment played the next biggest role. Conversely, factors that are associated with better health in early age, and that we hypothesized would decrease costs in later life (e.g. higher education, being white non-Hispanic), actually led to increased costs. The effect of these factors at older age is probably attenuated by a survival bias, with individuals more susceptible to the negative effects of low socio-economic status dying before they reach retirement age. However individuals of higher socio-economic status could also have greater access to and better knowledge of the healthcare resources and therefore use them more. This could explain our finding of lower expenditures among service workers, the group with highest proportion of blacks and with nearly the lowest mean education level, both associated with limited access to health care. However the survival bias hypothesis is supported by the smokers in our study having substantially lower expenditures than the non-smokers,
likely due to the “high-cost” smokers having already died by age 65. However, it is also possible that smokers are less attentive to their health needs and therefore less likely to seek medical treatments.

While chronic condition prevention efforts might be an effective measure towards reducing future costs in the population group that is still young, such preventive measures will not reduce the number of diseases or lower costs in the individuals who have already developed them. However, associated costs could be decreased by preventing disability and frailty associated with the illnesses.

When the number of chronic conditions is held constant, those older adults in our study who continued to work had substantially lower healthcare costs. These results imply that it might be the amount of disability associated with disease and not chronic disease itself that is responsible not only for older workers’ ability to work, but also for their reduced healthcare costs. In addition, the beneficial health effect of employment at older age via increased activity, social engagement, and access to better health care resources should also not be dismissed. If the latter factors are at play, then it may be that increasing the age of retirement in the US would also reduce health costs, even if Medicare eligibility remained at age 65.

Measures reducing disability associated with chronic disease could not only directly result in reduced costs, but could also enable more older individuals to seek employment. That in turn could result in further improvement in health and cost reduction, as well as in increased Social Security and Medicare contributions. The workplace may also provide a convenient venue for further disability reduction and health promotion measures, targeting both older and younger workers. In fact, workplace
chronic disease management programs might be more effective than lifestyle management interventions at reducing healthcare costs and hospital admissions in workers.\textsuperscript{191}

We found a trend (p < 0.10), towards increased expenditures and utilization among farmers when compared to white collar workers. Farmers are an occupational group with a high proportion of older workers, which suffers from limited access to care and low rates of health insurance coverage, potentially leading to increased costs upon Medicare eligibility.\textsuperscript{35,192-194} In addition, older farmers are often the ones working with heavy equipment, that is performing mostly sedentary jobs but with high risk of injury.\textsuperscript{31,35,195} A trend towards increased expenditures and substantially increased incidence of ER visits and number of hospital days calls for improved safety measures targeting this group, in addition to improved health care access among younger farm workers.\textsuperscript{31,195}

V. Limitations

This study used pooled cross-sectional self-reported data. Therefore strong causal inferences cannot be made, and inaccuracies in reporting are possible. In MEPS, only one person per household is interviewed and reports information for the entire household, which might result in bias and misreporting. The small number of farmers in the sample could result in our inability to detect any significant results in this group. Employment/occupation was measured during one week at the beginning of the year, and might not represent employment status during the rest of the year for which outcomes were measured. It was also unknown what the participant’s previous occupation history
was. Information on other risky behaviors such as alcohol consumption, as well as previous smoking history was not available. The chronic condition count did not include mood disorders, which are among the top conditions contributing towards increased healthcare spending. Standard model fit indices (i.e. RMSEA, CFI) could not be obtained with the MLR estimator used in this study, and standardized path coefficients could not be calculated.

VI. Conclusions

Our study found that factors traditionally associated with better health outcomes, such as higher socioeconomic status and non-smoking, resulted in increased health care costs at older age. In fact, health factors detrimental to health, such as smoking, actually resulted in substantially reduced costs in adults aged 65+, possibly due to a strong survivor bias. After controlling for the effect of socio-demographics, smoking status, and health insurance, it was the number of chronic health conditions and employment status that had the greatest effects on healthcare costs and utilization of the elderly population in our study. Measures reducing disability and frailty associated with chronic illness could not only result in decreased costs, but also enable sicker older individuals to seek or continue employment, resulting in further reduction in costs, increased incomes, and increased contributions to the Social Security and Medicare funds.
Chapter VI: Discussion and Conclusions

I. Overview

The proportion of older adults in the population is rapidly expanding, both in the United States as well as worldwide. The current generation of older adults are healthier than the previous generations, and as they are facing an increased number of years lived in retirement, an ever increasing proportion of the older population is choosing to remain employed past age 65. As a result, the effect of employment on health at older age has become a subject of great interest in the research community. However, the health-related quality of life (HRQL) and the healthcare expenditures and utilization patterns of older workers have not been well characterized. This research has examined a variety of health outcomes of US residents aged 65 years and older using the nationally representative pooled data from the NHIS (years 1997-2011) and MPES (years 2000-2009). Starting with examining the sociodemographic and health behavior predictors of several basic measures of health status in older adults (e.g. the number of chronic health conditions and functional limitations), this study then moved on to modeling the more complex consequences of these health states such as the health-related quality of life, health service utilization, and healthcare expenditures. We examined the effects of a variety of individual level predictors on these outcomes, including those that in the context of health service utilization causative pathways that are considered predisposing (e.g. education, occupation gender), enabling (e.g. income, health insurance coverage), and need (e.g. the number of chronic health conditions) factors.

With an increasing proportion of older adults choosing to work past retirement age, the effects of employment and particular occupations on their health was of
particular interest in this study. Furthermore, as the immediate effect of work was of interest, the occupation information used was for their current and not their lifetime longest-held occupations. Therefore nationally representative sources of data that contained information on older adults current employment status as well as their health outcomes of interest, were selected for this study.

We found that significant differences existed in the health outcomes of older workers and non-workers, while the differences between occupations of workers were relatively slight by comparison, potentially reflecting a strong selection bias for the working elderly overall compared to the rest of the elderly population (i.e. healthy worker effect). The differences in higher-level outcomes (i.e. HRQL, and healthcare utilization and expenditures) between workers and non-workers persisted even after controlling for the number of chronic health conditions, suggesting that at least some of the effect might be due to the beneficial effect of employment on health.

II. Elderly Population Characteristics

The results discussed in this section were pooled from both NHIS and MEPS analyses conducted for all three specific aims. Where relevant, and where these differ, both NHIS and MEPS estimates are given. While both surveys are designed to be representative of the US population, the years during which the data were collected differ, and therefore the estimates may differ slightly as well.
i. **Older Adults**

The average older adult represented by this study’s sample was between 74.4 and 74.6 years old (MEPS and NHIS estimates respectively), with females being on average a year older than males. The participants reported an average of 1.94 chronic conditions per person (MEPS), with over a half reporting previous diagnosis of hypertension and arthritis, and 61% reporting multiple (i.e. >1) condition (Table 6.1). Previous studies have found the prevalence of multimorbidity among older adults to be between 55% and 98%, depending on the conditions included.\textsuperscript{174} The majority of older adults were unemployed or retired, while the employed older adults were between 5-6 years younger than the unemployed (Table 5.1) and had fewer chronic health conditions.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>Any diagnosis</td>
<td>29,136</td>
<td>86.39</td>
</tr>
<tr>
<td>Two or more diagnoses</td>
<td>20,497</td>
<td>60.77</td>
</tr>
<tr>
<td>Arthritis</td>
<td>18,088</td>
<td>53.94</td>
</tr>
<tr>
<td>Asthma</td>
<td>3,102</td>
<td>9.21</td>
</tr>
<tr>
<td>Diabetes</td>
<td>7,155</td>
<td>21.25</td>
</tr>
<tr>
<td>Emphysema</td>
<td>1,816</td>
<td>5.40</td>
</tr>
<tr>
<td>Heart disease\textsuperscript{a}</td>
<td>10,746</td>
<td>32.01</td>
</tr>
<tr>
<td>Hypertension</td>
<td>21,725</td>
<td>64.64</td>
</tr>
<tr>
<td>Stroke</td>
<td>3,656</td>
<td>10.87</td>
</tr>
</tbody>
</table>

\textsuperscript{a}Heart disease included: myocardial infarction, angina, coronary heart disease, and other.
The majority of older adults (57.3%) had private insurance coverage, which in these analyses included any Medigap coverage as well as TRICARE. Table 6.2 illustrates predictors of private insurance coverage obtained using structural equation model described in chapter 5. Higher income and education, current employment, male gender, and being white non-Hispanic all increased the likelihood of having private insurance coverage. Some of these individuals are likely to have had employer coverage through their job.

Table 6.2: Predictors of Private Insurance Coverage as Compared to Public Coverage. Medical Expenditure Panel Survey.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Odds Ratios</th>
<th>CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female vs. Male</td>
<td>0.88</td>
<td>(0.82; 0.94)</td>
</tr>
<tr>
<td>Income category</td>
<td>1.38</td>
<td>(1.34; 1.42)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>0.98</td>
<td>(0.97; 0.99)</td>
</tr>
<tr>
<td>Black Non-Hispanic (vs White non-Hispanic)</td>
<td>0.57</td>
<td>(0.50; 0.65)</td>
</tr>
<tr>
<td>Hispanic (vs White non-Hispanic)</td>
<td>0.36</td>
<td>(0.30; 0.42)</td>
</tr>
<tr>
<td>Other race/ethnicity (vs White non-Hispanic)</td>
<td>0.40</td>
<td>(0.32; 0.51)</td>
</tr>
<tr>
<td>Chronic health condition (each additional)</td>
<td>0.97</td>
<td>(0.94; 1.00)</td>
</tr>
<tr>
<td>Education (years)</td>
<td>1.09</td>
<td>(1.08; 1.11)</td>
</tr>
<tr>
<td>Unemployed / Retired (vs White collar)</td>
<td>0.83</td>
<td>(0.73; 0.95)</td>
</tr>
<tr>
<td>Blue collar (vs White collar)</td>
<td>1.16</td>
<td>(0.93; 1.45)</td>
</tr>
<tr>
<td>Farmer (vs White collar)</td>
<td>0.79</td>
<td>(0.39; 1.59)</td>
</tr>
<tr>
<td>Service worker (vs White collar)</td>
<td>0.86</td>
<td>(0.68; 1.10)</td>
</tr>
</tbody>
</table>

There was a small percentage (0.4%) of uninsured individuals in the sample (Table 6.3). In our study, these individuals were younger, more likely to be female and non-white, have fewer chronic conditions, and be less educated. Employment was not associated insurance status; however, the highest proportion of the uninsured was
among farmers (2.3%). That is, the uninsured older adults represent the groups that either have not enrolled in Medicare yet, and are the youngest and healthiest proportion of the elderly; or have limited access and about it, or quite possibly are not eligible for Medicare coverage.\textsuperscript{15,196-198}

Table 6.3: Predictors of Having no Insurance Coverage as Compared to Public Coverage. Medical Expenditure Panel Survey.

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Odds Ratios</th>
<th>CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female vs. Male</td>
<td>1.64</td>
<td>(1.06; 2.53)</td>
</tr>
<tr>
<td>Income category</td>
<td>0.86</td>
<td>(0.73; 1.01)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>0.89</td>
<td>(0.85; 0.93)</td>
</tr>
<tr>
<td>Black Non-Hispanic (vs White non-Hispanic)</td>
<td>3.19</td>
<td>(1.43; 7.15)</td>
</tr>
<tr>
<td>Hispanic (vs White non-Hispanic)</td>
<td>9.02</td>
<td>(4.24; 19.19)</td>
</tr>
<tr>
<td>Other race/ethnicity (vs White non-Hispanic)</td>
<td>11.91</td>
<td>(5.12; 27.71)</td>
</tr>
<tr>
<td>Chronic health condition (each additional)</td>
<td>0.50</td>
<td>(0.39; 0.63)</td>
</tr>
<tr>
<td>Education (years)</td>
<td>0.86</td>
<td>(0.82; 0.92)</td>
</tr>
<tr>
<td>Unemployed / Retired (vs White collar)</td>
<td>1.00</td>
<td>(0.41; 2.43)</td>
</tr>
<tr>
<td>Blue collar (vs White collar)</td>
<td>1.77</td>
<td>(0.74; 4.25)</td>
</tr>
<tr>
<td>Farmer (vs White collar)</td>
<td>2.15</td>
<td>(0.28; 16.40)</td>
</tr>
<tr>
<td>Service worker (vs White collar)</td>
<td>0.62</td>
<td>(0.20; 1.95)</td>
</tr>
</tbody>
</table>

With regards to the HRQL scores, the US older adults had a mean PCS score of 42, and a mean MCS score of 52. As was mentioned in chapter 2, SF-12 component scores are scaled such that a score of 50 represents the mean of the general population with a standard deviation of 10. That is, on average, the participants of the study had physical functioning scores almost an entire standard
deviation lower than the general population, which is quite natural given that physical health deteriorates with age. However, older adults scored slightly higher on the mental health component of the SF-12 measure than the general population. A previous study of ambulatory Medicare beneficiaries found the mean scores for PCS and MCS to be 43 and 52 respectively. Therefore the PCS scores in our general older population are one point lower, on average, than those in an ambulatory older population.

The mean EQ-5D score in our study was 0.72. Several previous studies using 2000-2002 MEPS data found the mean HRQL of adults aged 60+ to be between 0.74 and 0.82. However all of these studies excluded from the analyses the EQ-5D data that was reported by proxies and not by the intended respondents. The excluded scores constituted approximately 12% of the sample in one of the studies, and the excluded participants were more likely to be black or Hispanic, not speak English, and have lower education and income – factors associated with lower EQ-5D scores. Our study included these scores completed by proxies in the analyses, and therefore provides more complete estimates of the actual reported EQ-5D scores; however it is subject to a bias due to proxy reporting of HRQL measures.

ii. Older Workers

The most common occupation group among those employed was white collar (7.9% of NHIS, and 10.67% of MEPS total sample). This finding is not surprising given that white collar occupations are generally physically less demanding and possibly more accommodating, and therefore might allow an older workers to
continue working longer. In addition, those older workers who were previously employed in more physically demanding jobs might transfer into white collar employment after they retire from their career employment. The proportion of men employed in blue collar and farming occupations was higher than that of women; there was a much higher proportion of women among service workers.

Among the unemployed/retired, 11.5% had incomes below the poverty level and 7.8% were near poor (Table 5.1). By contrast, among workers only 1-3% fell into either category, with over a half of white and blue collar workers earning high incomes (i.e. \(\geq 400\%\) of poverty level). This is possibly due to the worker’s incomes being increased by working. However, considering that the unemployed older adults were also older and sicker, as well as poorer, some older adults who live below the poverty level and might potentially like to increase their income by getting a job, might not be able to do so due to health reasons. The high earners among the white and blue collar workers might be healthier as a result of life-time higher incomes (workers with higher SES are less likely to change occupations past retirement age\(^{20}\)), and therefore better able to remain employed. They also might be more likely to work out of desire to remain engaged rather than out of necessity to make a living. The majority of the working poor were among the service workers, with 3.3% of older adults in this occupation being poor and 3.2% being near poor. Service jobs are low paying jobs requiring minimal skills. Therefore individuals employed in them are more likely to be those who work out of economic necessity, and who have trouble finding anything higher paying.
Approximately 5% of males and 22% of females reported having never worked (Table 2.1). Approximately half of male workers across all occupations had previously retired except for Service workers, among whom 64% of men had retired. On the other hand, the majority of working women have never retired. Women might be less inclined to retire due to lower incomes and greater financial need, and also they might be less likely to return to work after retirement.200.

Table 6.4 presents predictors of late life employment in each occupational group modeled as a part of the analyses in chapter 5. Workers in general were healthier, with physically demanding occupations being the healthiest. White collar workers were more highly educated. Service workers were more likely to be black females with low income levels, and most likely to be smokers. Farm workers were the healthiest group in terms of the number of chronic health conditions, least likely to smoke, and mostly male.

III. Older Adult Health Status – Specific Aim #1

In this specific aim, we characterized the health status of the US adult population aged 65 and over with respect to 4 health outcomes: multimorbidity (i.e. presence of multiple chronic conditions), presence of multiple functional limitations, fair/poor self-rated health, and low HALex score. The first three of these health status characteristics are sequential components of the causal pathway of HRQL, with HALex score being an HRQL measure. We examined the association between these outcomes and individual level demographic (i.e. age, gender, race/ethnicity), socio-economic (i.e. education,
Table 6.4: Predictors of Employment and Occupation in Later Life as Compared to Being Employed in a White Collar Occupation. MEPS Data.

<table>
<thead>
<tr>
<th></th>
<th>Odds Ratios</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Unemployed / Retired vs White Collar</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female (vs male)</td>
<td>1.61</td>
<td>(1.48; 1.74)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>1.16</td>
<td>(1.14; 1.17)</td>
</tr>
<tr>
<td>Black non-Hispanic (vs white)</td>
<td>0.94</td>
<td>(0.80; 1.10)</td>
</tr>
<tr>
<td>Hispanic (vs white)</td>
<td>0.82</td>
<td>(0.68; 0.99)</td>
</tr>
<tr>
<td>Other race/ethnicity (vs white)</td>
<td>1.36</td>
<td>(1.09; 1.70)</td>
</tr>
<tr>
<td>Education (years)</td>
<td>0.90</td>
<td>(0.88; 0.92)</td>
</tr>
<tr>
<td>Comorbidities (# of conditions)</td>
<td>1.25</td>
<td>(1.20; 1.30)</td>
</tr>
<tr>
<td>Smoker (vs non-smoker)</td>
<td>1.18</td>
<td>(1.01; 1.37)</td>
</tr>
<tr>
<td><strong>Blue Collar vs White Collar</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female (vs male)</td>
<td>0.12</td>
<td>(0.10; 0.16)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>0.86</td>
<td>(0.84; 0.88)</td>
</tr>
<tr>
<td>Black non-Hispanic (vs white)</td>
<td>1.16</td>
<td>(0.90; 1.51)</td>
</tr>
<tr>
<td>Hispanic (vs white)</td>
<td>0.79</td>
<td>(0.56; 1.10)</td>
</tr>
<tr>
<td>Other race/ethnicity (vs white)</td>
<td>0.59</td>
<td>(0.33; 1.06)</td>
</tr>
<tr>
<td>Education (years)</td>
<td>0.91</td>
<td>(0.89; 0.93)</td>
</tr>
<tr>
<td>Comorbidities (# conditions)</td>
<td>0.79</td>
<td>(0.73; 0.86)</td>
</tr>
<tr>
<td>Smoker (vs non-smoker)</td>
<td>0.92</td>
<td>(0.71; 1.19)</td>
</tr>
<tr>
<td><strong>Farm Worker vs White Collar</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female (vs male)</td>
<td>0.19</td>
<td>(0.08; 0.43)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>0.89</td>
<td>(0.84; 0.95)</td>
</tr>
<tr>
<td>Black non-Hispanic (vs white)</td>
<td>0.71</td>
<td>(0.28; 1.83)</td>
</tr>
<tr>
<td>Hispanic (vs white)</td>
<td>0.97</td>
<td>(0.43; 2.21)</td>
</tr>
<tr>
<td>Other race/ethnicity (vs white)</td>
<td>0.31</td>
<td>(0.07; 1.43)</td>
</tr>
<tr>
<td>Education (years)</td>
<td>0.84</td>
<td>(0.79; 0.89)</td>
</tr>
<tr>
<td>Comorbidities (# conditions)</td>
<td>0.79</td>
<td>(0.63; 1.00)</td>
</tr>
<tr>
<td>Smoker (vs non-smoker)</td>
<td>0.29</td>
<td>(0.13; 0.63)</td>
</tr>
<tr>
<td><strong>Service Worker vs White Collar</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female (vs male)</td>
<td>1.22</td>
<td>(1.01; 1.49)</td>
</tr>
</tbody>
</table>
Table 6.4 continued:

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>0.89</td>
<td>(0.88; 0.91)</td>
</tr>
<tr>
<td>Black non-Hispanic (vs white)</td>
<td>2.16</td>
<td>(1.68; 2.80)</td>
</tr>
<tr>
<td>Hispanic (vs white)</td>
<td>1.30</td>
<td>(0.93; 1.83)</td>
</tr>
<tr>
<td>Other race/ethnicity (vs white)</td>
<td>1.15</td>
<td>(0.75; 1.79)</td>
</tr>
<tr>
<td>Education (years)</td>
<td>0.93</td>
<td>(0.91; 0.95)</td>
</tr>
<tr>
<td>Comorbidities (# conditions)</td>
<td>0.84</td>
<td>(0.78; 0.91)</td>
</tr>
<tr>
<td>Smoker (vs non-smoker)</td>
<td>1.22</td>
<td>(0.94; 1.57)</td>
</tr>
</tbody>
</table>

employment/occupation), and health behavior (i.e. smoking and drinking status) characteristics.

As expected from the conceptual model and previous studies, individual level biological, behavioral, and social factors were all associated with the health outcomes in this study. Gender, education, race/ethnicity, age, and drinking/smoking history predicted all of the components of the HRQL causal pathway.\textsuperscript{135,142,146,147} As previously reported, higher education, employment, younger age, non-Hispanic white race/ethnicity, non-smoking, and current drinking were all associated with better outcomes on all health status measures. Male gender was associated with multimorbidity, although previous studies have found multiple conditions to be more prevalent among females.\textsuperscript{174} However, the chronic condition count variable in these NHIS-based analyses did not include arthritis, as this diagnosis was not assessed prior to year 2001. Arthritis is much more prevalent in women and was among the most prevalent conditions reported by MEPS participants.\textsuperscript{201,202}

The effect of gender and Hispanic race/ethnicity varied by outcome. Hispanics were more likely to report fair/poor health and poor (low) HALex score, but less likely to have multiple chronic conditions and functional limitations. As was discussed in chapter
3, this might be partially a result of limited access to healthcare in this group and consequently the under-diagnosis of chronic diseases. However, under-diagnosis would not explain lower odds of functional limitations, and therefore this result is also indicative of differential interpretation of what “good health” means in different cultures.124,153

Results also differed by outcome for men, who were more likely to report multimorbidity and fair/poor health, although were less likely to have multiple functional limitations or low HALex (a measure that included ADL limitations in its calculation). Thus, while being in poorer health, men were less likely to be limited due to their health. This could be explained by a lack of access to assistive technology among women, either for financial or other reasons; or due to cultural differences between men and women. Men are expected to be strong and not complain, and therefore they might be less likely to report limitations in activities. In addition, women are culturally more likely than men to do physically active housework, and therefore might be more likely to notice limitations in their activity level.

The differential findings for these two groups of older adults highlight the highly subjective nature of health status measures. Different items may not be measuring the same aspect of health; and different population groups use different criteria for assessing different aspects of their health. For example, when rating their health, some participants might report the specific health problems they have, while others might include their health behaviors, or focus on their general functioning. In addition, those of lower income and education level and with less healthy lifestyles reportedly tend to be more optimistic about their health and rate it higher. The effect of this differential reporting may influence the HALex assessment, as self-rated health is a component in its
calculation. On the other hand, differential results on functional limitations and comorbidity might indicate issues in access to care or being able to afford assistive technologies. The use of four different health status measures in this study allowed a more complete evaluation of health.

The differential results for men also highlight the fact that chronic conditions do not necessarily result in functional limitations if the conditions are well managed and if the person has proper access to assistive technology.\textsuperscript{110,137,156,157} In fact, morbidity, disability, and frailty in older adults are distinct and relatively independent phenomena that overlap only partially.\textsuperscript{110,206} For example, in a study by Fried et al., only about 11\% of older adults with multiple chronic conditions also reported having a disability.\textsuperscript{206} The implication of this finding is that productive life can be extended into later years even in the presence of chronic disease.

The effects of smoking on health varied by outcome more subtly. The odds of fair/poor health, functional limitations, and low HALex where higher in current smokers vs. former smokers when compared to never smokers. However, for multimorbidity this pattern was reversed: the former smokers had higher odds of multimorbidity than current smokers when compared to never-smokers. This finding suggests selective survival, as smokers who were susceptible to the harmful effects of smoking and who would develop multiple chronic diseases have already died, while those who quit smoking have still developed a greater number of chronic illnesses, but not enough to die from them just yet.\textsuperscript{187,207,208} Some current smokers might also be light smokers and therefore not suffering from the effects of smoking as much.\textsuperscript{208} Regardless, current smoking resulted in
poorer health outcomes than former smoking on other measures, even in the absence of an increase in the number of chronic diseases.

With respect to alcohol consumption, current drinkers had the best health outcomes, with former drinkers faring the worst. This is consistent with multiple previous studies in both general population and older adults. In older adults, current drinking has been linked to better mobility and mental function, lower mortality, and higher HRQL.\textsuperscript{148-150,152} The simple explanation for this finding is the poorer overall health of former drinkers, which causes them to stop drinking.\textsuperscript{207} Regular alcohol consumption in general population has also been linked to lower healthcare utilization, believed to be partially due to better health (and partially to delaying care in heavy drinkers).\textsuperscript{209}

Employment was associated with improved health outcomes across all measures. The more physically demanding occupations (i.e. non white collar) had better health outcomes than white collar workers, suggestive of a stronger healthy worker effect present in these more physically demanding occupations.\textsuperscript{145} However, work at older age can also be beneficial for health by the mechanisms of increased social engagement, increased physical activity, and increased income and access to better health insurance.\textsuperscript{12,13,50}

It is appropriate to bring up the differential item functioning here again. As mentioned above, individuals might interpret the same questions differently and thereby respond differently. For example, older farmers are more likely to assess their health status based on whether they are able to continue working rather than the presence of chronic illness.\textsuperscript{35} Therefore, to an extent, these individuals’ favorable rating of their own health might be a consequence of their very ability to still work, and therefore in effect be
“caused” by employment, however not for the reasons commonly discussed. In fact, in occupational context, the individual’s rating of health might even depend on how much they like their job and how much they want to keep on doing it. Based on these factors an individual might either play up or play down their self-report of health measures. This of course does not mean to imply that the measures used are not valid to measure differences between occupations. Self-rated health in particular has correlated well with the individual’s health status determined medically. However, self-reported subjective measures of health should be interpreted in combination with other measures that cover alternative aspects of health status.

IV. Older Adult Health-Related Quality of Life – Specific Aim #2

In this specific aim, we used structural equation modeling to predict three distinct measures of HRQL in older adults: PCS (physical health component summary of SF-12), MCS (mental health component summary of SF-12) and EQ-5D. The model included individual level predictors that were treated as exogenous variables (i.e. education, age, smoking status, gender, race/ethnicity) and individual level predictors that were treated as mediators (number of chronic conditions, income, health insurance, employment/occupation).

H2: We hypothesized that workers in white collar occupations who continued working past retirement age would demonstrate higher quality of life relative to retired peers, while workers in other occupations would have lower quality of life relative to their retired peers after adjustment for other factors. However, we found that all occupational groups generally had better HRQL scores than the unemployed. Blue collar
workers had higher MCS scores, and farm and service workers had better EQ-5D scores
than the white collar workers, however no differences in PCS scores were detected.
Therefore, older workers of all occupations, after controlling for the number of chronic
health conditions, were in approximately similar and relatively good physical health.

Table 6.5: Multivariable Logistic Regression Results Modeling Multiple Functional
Limitations (NHIS data).

<table>
<thead>
<tr>
<th></th>
<th>Odds Ratio*</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unemployed / retired vs white collar</td>
<td>1.86</td>
<td>1.73 – 2.00</td>
</tr>
<tr>
<td>Service vs white collar</td>
<td>0.83</td>
<td>0.72 – 0.95</td>
</tr>
<tr>
<td>Farm vs white collar</td>
<td>0.96</td>
<td>0.71 – 1.30</td>
</tr>
<tr>
<td>Blue collar vs white collar</td>
<td>0.85</td>
<td>0.73 – 0.99</td>
</tr>
</tbody>
</table>

*Results were adjusted for gender, education, race/ethnicity, age, smoking and
drinking status, and survey year

This point is important because due to variability of impact that different health
conditions included in the condition count variable might have on health, this count
variable might seem like not the most appropriate way to control for the healthy worker
effect. However, it has eliminated the physical health status differences between the
worker groups, even though differences in disability rates exist between occupations
(Table 6.5). There were still differences in the mental health scores, however we were not
able to include mental health measures in the chronic disease count as these were not
measured in MEPS.
This homogeneity in the physical health of older workers of different occupations also indicates that there might be a baseline level of good physical health that predisposes a worker to be working at older age. That is, the older person needs to be at least physically mobile and mentally agile enough to make it to work. However, this threshold level of health does not seem to be any higher for somebody who does physical work than for somebody who works in a white collar occupation. On the other hand, such homogeneity of physical health status across occupations suggests that whatever the physical health benefits from employment at older age are, they do not depend as much on the kind of work the person does as on the very act of doing it.

Workers in the physical occupations did better on HRQL scores, which contained a mental health component. The potential mental health benefits of employment could stem from social engagement associated with employment, feeling useful, as well as increased income, as the mediation pathway results in our study indicate. This benefit might be lower for white collar workers, the most educated of all occupational groups, who are more likely to have better opportunities to stay engaged outside of the workplace than other workers, and therefore might not benefit as much from increased social engagement associated with work. More educated workers are also more likely to have been in the same occupation prior to retirement age, thereby benefiting from higher incomes than their peers throughout their lives, and therefore less likely to reap the mental health benefits of income increase due to employment at older age.

Employment in general, regardless of occupation, was a strong predictor of better scores on all three measures, and it was the next strongest predictor of PCS after the number of chronic conditions. The effect of employment was partially mediated by
income. However the mediation effect was stronger for MCS and EQ-5D than for PCS relative to the direct effects; and most of the effect on PCS was direct. Even after adjustment for comorbidities, there might still be a healthy worker effect as discussed above, with a certain threshold level of physical and mental health required to work at older age. However, our data also suggest that beneficial physical and mental effects of employment are present.

A previous study by Caban-Martinez at al. found the EQ5D scores to be higher in white collar workers than in other groups.\textsuperscript{178} That study used NHIS data linked to MEPS, which by design assessed EQ5D scores 1-2 years after the employment/occupation status was assessed. Our results indicate however that HRQL scores are higher in more physically demanding occupations. It is possible that the HRQL scores decline more rapidly in those employed in more physically demanding occupations, resulting in the higher scores of white collar workers a year or two later. Older workers in more physically demanding jobs might also be more compelled to retire earlier. Retirement is not universally associated with negative health outcomes, however in individuals of lower socio-economic status (SES) it is more likely to lead to declines in health.\textsuperscript{48,49} The differences in health consequences of retirement between higher and lower SES workers might be contributing to the previously observed higher HRQL scores in white collar workers at 1-2 year follow-up.

\textbf{H\textsubscript{2-2}}: We also hypothesized that while older workers in race/ethnic minority subgroups would experience lower quality of life than non-Hispanic white older workers, the direct effects of race/ethnicity would be eliminated by introducing a mediation effect via occupation. We found that only the MCS scores were affected by race/ethnicity.
Individuals in the non-Hispanic black and other categories did in fact have lower MCS scores; however, the direct effect of race/ethnicity was only slightly reduced by introducing a pathway via employment. A previous study by Fleishman and Lawrence also found race/ethnicity to be associated with MCS scores only and not with PCS scores.\textsuperscript{169} This study of the general adult population found blacks to have higher MCS scores than whites and Hispanics; however, they found that the black-white differences were due to MCS differential item functioning (i.e. different interpretation of questions by people in different subgroups) in these populations rather than a true difference. Differential item functioning within the MCS measure across racial/ethnic groups might explain our absence of findings for race/ethnic groups on any other HRQL measure besides MCS. We did not find a difference between blacks and whites, possibly because the effect was eliminated by controlling for socio-economic status characteristics and chronic conditions. In a previous study of older adults based on the Chicago Health and Aging Project data, controlling for these factors as well as the cognitive function eliminated black-white differences in the HRQL scores.\textsuperscript{142}

The number of chronic health conditions was associated with the lowest scores in all of the HRQL outcomes. Chronic conditions explained the largest amount of variance in PCS and less variance in MCS, with most of the effects being direct. Many previous studies have found a similar association. For example, both physical and mental health components of SF-36v2 were negatively associated with the number of chronic conditions in a sample of Medicare beneficiaries; and a Swedish population-based study found the number of chronic conditions to be the strongest predictor of HRQL scores.\textsuperscript{72} Brettschneider et al. also found that an increasing number of conditions in older adults
led to an increasingly poorer HRQL scores; and they identified depression, obesity, and Parkinson’s disease as the conditions with the greatest impact on HRQL.\textsuperscript{211} We did not have information on these conditions available in our analyses; potentially the effects of chronic illness in our study could have been stronger if these were included in the condition count. The number of chronic conditions also significantly mediated many of the indirect effects in the model, especially in predicting PCS.

Current smoking was associated with a negative change on all HRQL scores, but the change was especially pronounced by comparison to other predictors for MCS. Mental health scores were also especially affected by current smoking in a study of patients with heart disease.\textsuperscript{212} Smoking was associated with a slight increase in PCS scores via its effect on the number of chronic conditions. Unlike in the analyses based on NHIS data, smoking was associated with a decreased number of chronic conditions in MEPS analyses (Table 6.6). The MEPS did not distinguish between former and never smokers, and therefore people in these two categories were grouped together. However, as NHIS results discussed above indicate, in the older adult population never-smokers have the lowest number of chronic conditions, and former smokers have the highest number of conditions, even higher than current smokers. By grouping these two together, the “non-smoker” category in MEPS ended up with more conditions than the “current smoker.”

V. Older Adult Healthcare Expenditures and Utilization - Specific Aim #3

In this specific aim, we used SEM to model total healthcare expenditures as well as the annual number of ER visits and number of hospital nights. We used the same
Table 6.6: Multivariable Linear Regression Predicting the Number of Chronic Health Conditions. MEPS data.

<table>
<thead>
<tr>
<th></th>
<th>Effect</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>0.05</td>
<td>(0.01; 0.09)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>0.02</td>
<td>(0.02; 0.02)</td>
</tr>
<tr>
<td>Black Non-Hispanic (vs White)</td>
<td>0.23</td>
<td>(0.16; 0.30)</td>
</tr>
<tr>
<td>Hispanic (vs White)</td>
<td>-0.19</td>
<td>(-0.28; -0.10)</td>
</tr>
<tr>
<td>Other race/ethnicity (vs White)</td>
<td>-0.13</td>
<td>(-0.24; -0.02)</td>
</tr>
<tr>
<td>Education (years)</td>
<td>-0.04</td>
<td>(-0.05; -0.03)</td>
</tr>
<tr>
<td>Smoker (vs Non-smoker)</td>
<td>-0.10</td>
<td>(-0.17; -0.03)</td>
</tr>
</tbody>
</table>

individual-level predictors for this model as for modeling HRQL.

$H_{3.1}$: We hypothesized that the total healthcare expenditures would be significantly lower in older workers versus older non-workers, even after adjustment for sociodemographic factors and co-morbidities. Our results confirmed this hypothesis.

Employment was among the strongest predictors of reduced healthcare expenditures and healthcare use after controlling for all other factors. Previous studies have found that health service use does not increase substantially after retirement beyond what is explained by need, and therefore is not associated with work-retirement transition per se. However, in our cross-sectional study, there was a substantial difference between workers and non-workers, even after controlling for health conditions and sociodemographic factors that would affect access to care. This suggests an actual difference in utilization between the groups rather than a change that happens in transition from
work to non-work. Referring back to the original conceptual model of healthcare utilization by Andersen and Newman discussed in chapter 1, the predisposing, enabling, and need factors contributing to health service use might differ between workers and non-workers beyond what was controlled for in this study. For example, as indicated by our results for HRQL measures above, differences in physical and mental health functioning between these groups persisted even after controlling for the number of health conditions.

Another implication of the retirement event not contributing substantially to the amount of healthcare utilization, and of actual employment status in our study contributing substantially to it, is the strong effect of work, in and of itself, on the health outcomes. The lack of differences between occupations indicates that, among older workers, the type of work performed does not matter as much as the very act of engaging in it.

**H3-2**: We also hypothesized that the total healthcare expenditures would be significantly higher in blue collar than in white collar older workers, and this relationship would be only partially mediated by the type of health insurance coverage. However, we did not detect any expenditure differences between blue collar and white collar workers. The only difference between occupations we found was lower healthcare expenditures in service workers.

We found the number of chronic conditions to be the strongest predictor of all outcomes. Health status as measured by number of chronic conditions has previously been identified as a strong driver toward healthcare use, even a stronger determinant of it than financial factors. In fact, some studies suggest an almost exponential relationship between the number of conditions and healthcare expenditures and use. Again, our
chronic condition measure was limited by the information available in the data, and therefore might not have included some conditions that contribute to expenditures to a large extent, such as depression.86

Current smoking was associated with half the odds of having an expenditure as well as with a lower amount of expenditure, which could be due to a survivor bias of these older smokers. However, there also might be misclassification bias in our data due to how the variable was coded in MEPS, as described above. The category of non-smokers included former smokers, who in NHIS sample were sicker than the current smokers among older adults. However, there is some truth to lower expenditures among older smokers.186,188 It was previously predicted that smoking cessation would lead to increased lifetime healthcare spending, and progress toward a smoke-free society would ultimately lead to an increase in costs in the long run because the premature deaths would be prevented and people would live long enough to incur healthcare costs. Former smokers have especially high costs in the year after they quit smoking.213 A study of Bank of America retirees found smoking to be associated with lower healthcare costs214, however, its participants were younger, relatively well educated middle class individuals who were not disabled or unemployed prior to retirement age, representing a homogenous sample of relatively healthy people. Smokers’ mortality continues to be high relative to non-smokers into old age, with heavier smoking being associated with greater mortality.207,208

Smokers might utilize healthcare less because they delay seeking it,215,216 and therefore some of the reduced expenditures in smokers in our study might be due to older individuals not seeking timely care. In addition, poor health behaviors tend to correlate,
e.g. smokers are also more likely to be heavy drinkers and have high BMI, and each of these conditions is associated with a decrease in preventive service use. On the other hand the decrease in healthcare utilization associated with each of these poor health behaviors might have a different pattern, with smokers being less likely to use preventive care, and heavy drinkers being less likely to use routine care (i.e. associated with illness). Therefore in an individual who smokes and/or drinks might also have other unhealthy behaviors that were not controlled for in this analyses, and the effects of these on healthcare use might have additive effects utilization and costs. In this study we were not able to control for the effect of other behaviors, and all of the effects could have been attributed to smoking. Therefore, the estimated effect of smoking could be an over- or an underestimate of its actual effect.

A recent retrospective cohort study by Moriarty et al. found smoking to be associated with increased costs in retirees, and the costs were decreased by controlling for the number of chronic conditions. The results of the study were controlled for BMI, although not for any other potential confounders. In addition, the sample of the retirees was younger than our sample, it included individuals younger than 65 years (25th percentile age was 62 years), and came from a sample with continuous insurance coverage over the duration of follow-up. Another study found that in a sample of young-old relatively healthy adults, smoking was associated with increased costs, however this study also did not control for potential confounders. We found that in an older sample, and after adjustment for not only the effect of chronic conditions, but also other factors affecting costs, the direct effect of smoking vs. former and never smoking was actually to decrease costs.
Older service workers had the lowest healthcare expenditures of all occupations. While service workers had better health outcomes on the other outcomes examined previously, such as EQ-5D scores and multiple functional limitations, this occupational group was also poorer, less educated, more likely to be black and female, and more likely to smoke than other occupations. That is, they were likely to be coming from a disadvantaged background and possibly experiencing access to care issues and habitual underutilization of care, in addition to poor health behaviors. Older adults in population subgroups that traditionally have access to healthcare issues, such as black older adults, continue to underutilize healthcare even once they have insurance coverage and adequate access to care, and service workers of pre-retirement age are likely to experience difficulties with access to preventive care. However, considering better outcomes on other health characteristics as well, there might also be a survivor bias. Same as discussed with smoking above, the individuals who were susceptible to the negative effect of lower socio-economic status might have either died already or were sick to the point of no longer being able to work. The survivor bias is also consistent with our finding of higher costs being associated with the factors generally linked to better health outcomes, such as higher education and income and non-Hispanic white race/ethnicity as compared to black. However, from the access to care perspective, the more educated and privileged individuals might also have access to more healthcare resources and use them more as a result.

There was a trend towards higher expenditures and healthcare utilization among farmers. Farmers are a relatively small occupational group, however older adults make up a large proportion of this workforce, with approximately one third of American farms
being operated by workers aged over 65 years. The trend estimate of increased use was especially high for ER visits. A previous review of ER utilization by older adults showed that factors that improved access to primary care reduced ER utilization; and maybe this is what the farmers need – better access to primary care. This can be done by addressing predisposing (e.g. health beliefs about regular health maintenance through health education) and enabling (better availability and access to physician services at a regular source of care) factors; however it is essential to address the need based factors, such as the actual health status. Our results indicated that farmers are in better health than most older adults and have the lowest smoking rates. Although the decreased number of chronic illnesses might be due to difficulties in access to care and traditionally high proportion of the uninsured in this group resulting in underdiagnoses, farmers’ increased use of ER health services is also likely to be associated with occupational injury. Among farmers, older individuals are most likely to perform jobs associated with operating heavy machinery, and are common victims of severe injuries resulting from such mechanisms as tractor overturns. Therefore the finding of potentially increased ER visits among them calls for improved safety measures associated with machinery in farm work.

VI. Conceptual Model Revisited

We found that the conceptual model proposed in chapter 1 predicted our outcomes of interest well. All the pathways tested in the prediction of each group of outcomes were confirmed. The effect of race/ethnicity on HRQL outcomes was only present for MCS using the MEPS data; however different race/ethnic groups varied in outcomes on HALex, and therefore the pathway between race/ethnicity and HRQL
measures should be retained. In the future studies, BMI should also be included in the model as an additional health behavior parameter. Our analyses were adjusted for the effect of household income, however in older adults income might not constitute the best representation of one’s financial status. Therefore we suggest including an assessment of wealth, using measures such as home ownership, in future studies. Based on previous reports, in future studies we also suggest adding a pathway from HRQL to healthcare expenditures and utilization, as presented in figure 6.1.

We did not find substantial differences in health outcomes between occupations. However, there is plenty of previous evidence to the effect of retirement on health. Previous retirement history, and therefore motivation for work, might therefore play a greater role in affecting health outcomes at older age than the nature of the work itself. Therefore we suggest in future studies including previous retirement information in combination with current employment instead of occupation. We did not include retirement information in these analyses in order to preserve sufficient sample sizes within occupations. However in future studies careful attention of retirement status and movement in and out of the workforce, as well as motivation for such movement, is essential.

The number of chronic disease diagnoses was included as a predictor of health outcomes as well as employment/occupation in all our SEM models, partially with the intention to control for the healthy worker effect. However, as not all chronic conditions have an equal impact on the person’s functioning, the number of functional limitations might be a better measure of an individual’s ability to work, the kind of work one is able to do, as well as one’s propensity to retire. Like the chronic health conditions, functional
limitations are a part of the HRQL causal pathway and demonstrated a similar distribution across occupations (Tables 6.5 vs. 6.4). However, presence of functional limitations does not depend on access to care, as no diagnosis needs to be made, and it might provide a better measure of the activities one is able to perform, such as work. In addition, in MEPS functional limitations were assessed at the beginning of the year, versus the end of the year for chronic disease diagnoses. Therefore functional limitations would provide a more accurate measure of health status during the year when modeling annual healthcare expenditures. We suggest including the number of functional limitations
limitations in future studies instead of or in addition to the number of chronic conditions, as illustrated in Figure 6.1.

VII. Strengths and limitations

i. Limitations

An important characteristic of the data used in this study was that they were collected cross-sectionally; therefore the results of the analysis cannot have any causal implications. While MEPS is a panel survey which contains 2 years of longitudinally collected data, the relationship between worker’s current occupation/employment status and the outcome variables at or around the time of employment was of interest in this study, and therefore these data were treated as cross-sectional.

A major limitation to the use of the NHIS and MEPS data are the self-reporting of data without objective confirmation, which can result in both under-reporting and over-reporting by the participants widely varying by chronic disease category. In addition, reports are obtained from a proxy in some cases in the MEPS. For example, in MEPS one adult reports information for the entire household, which can lead to inaccurate reporting. Another issue was that the annual response rates to the 1997-2011 adult sample person interview averaged 70.3% (range: 60.8%-80.4%). MEPS response rates ranged from 56.9% in 2007 to 70.7% in 1996, and averaged 63.5%. The household and individual weights are designed by the NHIS and MEPS to take into account these non-response rates, and all the analyses were adjusted for these weights.
In this study, we did not have the capacity to control for the household financial assets. However, these in addition to income and education, play a major role in health even at older age. Furthermore, as age increases, wealth, or financial assets, become a more important predictor of health than income or education.

Determination of current employment status in the NHIS and MEPS may not represent the longest or even most important occupation from health effects point of view. However, employment at old age was specifically of interest in this study, and therefore the occupation data at that time point was deemed most appropriate for the study. Although concurrent employment information was appropriate for use in this study, lifetime occupational exposures might affect HRQL in older adults. For example, shift work earlier in work life can affect sleep quality, and as a result might affect HRQL after retirement. However, information about such exposures was not available in the data used. The employment/occupation information used in this study also did not include prior retirement information, which could be an important predictor of health outcomes. In MEPS, employment information used for annual healthcare utilization and expenditure analyses was collected at the beginning of the year and might not have represented the accurate employment status during the majority of the year for which outcomes were measured.

The chronic condition count variable included a variety of conditions, from asthma to stroke to angina. These do not have an equal effect on disability, HRQL, or healthcare use, but they were nevertheless treated equally in this study. The distributions of different conditions across subgroups of older adults, such as
occupations, might vary leading to inaccuracies in the estimation of the effects of chronic disease diagnoses on outcomes.

In this study, the number of chronic conditions in MEPS analyses was used as a proxy measure of health status that would potentially eliminate the healthy worker effect, as it was identified as one of the measures relevant to labor force participation. However, due to it being unable to account for the varying amounts of disability associated with specific illnesses, a different measure of disability (e.g., functional limitations) might be better suited for this purpose in future studies. In addition, we were not able to include some of the diseases that can potentially have a great impact on all outcomes of interest, as well as on the ability of an older individual to work, e.g. depression, osteoporosis, and obesity for all specific aims, as well as arthritis for specific aim #1.

In MEPS, the chronic disease diagnosis was measured at the end of the survey year for which the outcomes were measured. That is, for some individuals who were diagnosed with the disease at the end of the year, the expenditures and utilization results might not reflect the actual expenditures and utilization associated with chronic disease management and care.

In the calculation of the HALex index, one of the limitations assessed was the inability to work due to health reasons. This could potentially lead to an overestimation of the association between employment and HALex. However, this activity limitation was only a minor component of HALex calculation, and probably only affected the results to a small degree.
In the MEPS, information about other health behaviors (such as drinking) was not collected, and smoking information was only classified as “yes/no.” The consequences of such crude classification are discussed above in subsection IV of this chapter.

Because the MLR estimator was used for SEM modeling, standard model fit indices (such as RMSEA, CFI, etc.) could not be obtained. The WLS estimator was used on a subsample of the data to obtain the fit indices. However, the WLS estimates did not match those estimated using MLR, and therefore the model fit information could not be used. Mediation was assessed for specific aim #2, however indirect path estimates are likely to be inaccurate as categorical mediators had to be treated as continuous variables; and because cross-sectional data does not really allow making causal inferences.

We included HRQL scores reported by proxies in this study, however many previous studies have excluded them.67,167,199 The scores reported by proxies are more likely to belong to ethnic minorities; and there are systematic differences in HRQL scores reported by ethnic minorities who speak or do not speak English well.239 Therefore, there might be some bias associated with misreporting by these groups; however it allowed the inclusion of more complete data in the analyses.

ii. **Strengths**

Despite the limitations, the NHIS and MEPS pooled data represent a unique view of the risks and causes of health outcomes for the US older adults over a decade and a half (NHIS 1997-2011 and MEPS 1996-2009). Both datasets are nationally
representative, and are ideally suited to examine older worker health. The use of the two data sets in combination provided a comprehensive and cost-effective assessment of the US older adult health population. By their very design, the NHIS and MEPS datasets capture the diversity of the older US worker population (e.g., race/ethnicity, educational attainment). The large number of older workers included in these databases is a significant advantage combined with the probability sampling of the whole US population. Prior studies have suffered from the lack of data for the whole aging US civilian workforce, indirect and possibly inaccurate report of occupation, the lack of population based prevalence rates, and the lack of data on women and race-ethnic subpopulations. The NHIS and MEPS data and more sophisticated statistical analyses allowed for the exploration of the morbidity differences between occupations for US older workers and other subpopulations.

The major strength of MEPS data is that self-reported healthcare utilization data are verified and supplemented through the healthcare provider. This provides more accurate and complete information on utilization and costs that could be obtained by self-report alone.

A strength of the study was the use of the multiple outcome measures for each outcome of interest. This allowed for detection of differential effects that might have otherwise been ignored. The availability of some MEPS variables at several points during the survey year provided an opportunity to select a time point at which most relevant information was measured. For example, employment/occupation information was available at three different times during the year, and two of these time points were selected for use in our study. In modeling HRQL outcomes, the end-
of-the-year data on chronic condition diagnosis allowed us to include in the analyses conditions that might not have been diagnosed yet at the time of HRQL assessment, but that were already potentially affecting the person’s functioning.

Finally, a major strength of the study is in the complex statistical methods used for modeling, namely SEM. This set of techniques allowed to model all the relationships of interest simultaneously, while also controlling for indirect effects. To our knowledge, this is the first study to use SEM techniques for modeling HRQL and healthcare expenditure outcomes in a sample of older US workers.

VIII. Policy Recommendations

The findings of this study indicate that regardless of the type of work performed, older adults might benefit from the very act of being employed. Furthermore, the mental health benefits might be the greatest for those who actually stand to improve their income by working (and who constitute a large proportion of the currently unemployed older adults). Therefore, opportunities should be created for easier employment for those older adults who wish to work. The most effective approach to creating such opportunities would address both the older adults’ willingness to work, as well as the employers’ willingness to hire and retain older workers.

There are a number strategies which could be implemented to encourage older adults to seek employment, including changing the Social Security provisions so that the older workers are not “penalized” with reduced benefits for working after retirement. Employers and older workers may also benefit by making a variety of part-time, flexible, and telecommuting work options available.6 This could not only ease the demands on
older workers’ health, but also allow them to ease into retirement, potentially avoiding the negative consequences of an abrupt transition to retirement. The development and funding of government training programs that would train older adults in the skills necessary in the current job market could be created with a particular focus on the less educated and lower-skilled older workers. Finally, health promotion programs targeting older workers should focus on chronic disease management and disability prevention, as well as increased physical activity, in order to help older adults to achieve and maintain a basic level of good health necessary for work.

There are also many policies options for encouraging employers to hire older workers. First, the employment compensation system should be transitioned from one that is seniority-based to a performance-based system. This would allow employers to hire older employers at a reduced cost without having to pay extra for the added experience. This policy would have the added benefit of reducing ageism in the workplace due to inequality in pay for workers of different ages. Medicare should be mandated as the primary payer for those older adults who are covered by employer-provided insurance. This will substantially reduce employer healthcare costs, and help reduce the cost-associated reservations on hiring older workers. Older workers should be made exempt from the Social Security payroll tax to which employers must contribute. Employer tax credits should be provided for the training of older workers, as well as for the implementation of accommodations targeting the needs of older adults in the workplace. Tax credits could also be issued for to the reorganization of the workplace that would reallocate more physically demanding tasks to younger workers. Similarly, tax credits could be structured to encourage employers to provide the same training
opportunities to their older employees as they would to younger ones. This could be especially useful to older workers transferring from a career job to non-career employment.

Workplace wellness programs targeting older employees and promoting such activities as walking and yoga could decrease healthcare costs of older workers and be cost-saving for employers. Encouraging physical activity in the workplace could also lead to decreased rates of severe injury among older workers due to their better physical shape, further decreasing employer, worker and societal costs. Finally, for older adults to continue being productively engaged into older age and benefiting from it, a change in attitudes is needed to reduce ageism in the workplace and help employers see older worker as an asset, productive and knowledgeable, rather than a burden. For that purpose, educational materials can be developed and distributed to employers that would promote awareness about the value and productivity of older workers.

**IX. Future Research Directions**

The cross sectional data used in this study did not allow exploration of the direction of the relationship between employment and health, though previous studies clearly indicate it to be bidirectional. Health outcomes as predicted by employment/occupation might have different patterns of change over time across different occupations, as was discussed with regards to the EQ-5D results above. Therefore, the long-term effects of employment in a specific occupation should be explored further in order to identify the groups that could best benefit from certain types of employment in the long term. As a 2-year panel survey, MEPS offers 1 year of follow-
up data with 2 measurements of HRQL and 5 measurements of employment/occupation status for most adult participants, and it can be used as a starting point for exploring this longitudinal relationship. Other sources of data, such as the Health and Retirement Study (http://hrsonline.isr.umich.edu/), can be useful for this purpose as well.

We found that current smoking was associated with poorer health outcomes in NHIS analyses; however, it was associated with lower healthcare expenditures and utilization in the MEPS. Several previous studies of healthy older adults have found smoking to be associated with increased expenditures. Those studies did not use representative samples of the older population; however they bring to light the complexity of the relationship between smoking, health outcomes, and healthcare costs. MEPS data is perfectly suited for examination of expenditures; however it does not contain detailed enough smoking information. However, MEPS data can be linked to the NHIS for a subsample of older individuals, and the more precise smoking information could be obtained through NHIS allowing for better exploration of this relationship.

The results of our study also indicated that cross-sectionally the differences between older workers of different occupations are minor. Previous studies however have linked retirement status to major changes in health outcomes, and we suggest including it in future models. In subsection VI above we suggested some additional modifications to the conceptual model, such as using the functional limitations count variable instead of the chronic health condition count used in this study to adjust for the healthy worker effect. The conceptual model presented in Figure 6.1. should be used as the framework for future studies exploring the association between employment/retirement status and HRQL and healthcare expenditures. In addition, MEPS data in the context of SEM will
allow for simultaneous modeling of both outcomes within the same model, and will allow to explore the mediational pathways of the effect of employment/retirement at older age on healthcare expenditures and utilization via HRQL. In addition, trends in the outcomes of this study over the years of data available should be examined. Ultimately, we still need to understand the healthy worker effect among older adults better, as well as to identify the specific health benefits associated with different occupations.
References


5. Munnell AHS, Aubry, J.P. Employer survey: 1 of 4 boomers won't retire because they can't. Center for Retirement Research at Boston College; December 2006.


35. Reed DB, Rayens MK, Conley CK, Westneat S, Adkins SM. Farm elders define health as the ability to work. Workplace health & safety. Aug;60(8):345-351.


77. Fronstin P, Salisbury DL, VanDerhei JL. *Funding savings needed for health expenses for persons eligible for Medicare*.


Appendix A: NHIS and MEPS Variables Used for Modeling

I. NHIS

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Type</th>
<th>Length</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AASMEV</td>
<td>Num</td>
<td>3</td>
<td>Ever been told you had asthma</td>
</tr>
<tr>
<td>AGE_P</td>
<td>Num</td>
<td>3</td>
<td>Age</td>
</tr>
<tr>
<td>ALC5UPYR</td>
<td>Num</td>
<td>4</td>
<td>Number of days had 5+ drinks past year</td>
</tr>
<tr>
<td>ALCSTAT1</td>
<td>Num</td>
<td>3</td>
<td>Alcohol Drinking Status: Recode</td>
</tr>
<tr>
<td>ANGEV</td>
<td>Num</td>
<td>3</td>
<td>Ever been told you had angina pectoris</td>
</tr>
<tr>
<td>CANEV</td>
<td>Num</td>
<td>3</td>
<td>Ever told by a doctor you had cancer</td>
</tr>
<tr>
<td>CHDEV</td>
<td>Num</td>
<td>3</td>
<td>Ever told you had coronary heart disease</td>
</tr>
<tr>
<td>DIBEV</td>
<td>Num</td>
<td>3</td>
<td>Ever been told you had diabetes</td>
</tr>
<tr>
<td>DOINGGLW</td>
<td>Num</td>
<td>3</td>
<td>What was -- doing last week?</td>
</tr>
<tr>
<td>EDUC</td>
<td>Num</td>
<td>3</td>
<td>Highest level of school completed</td>
</tr>
<tr>
<td>EPHEV</td>
<td>Num</td>
<td>3</td>
<td>Ever been told you had emphysema</td>
</tr>
<tr>
<td>FLA1AR</td>
<td>Num</td>
<td>3</td>
<td>Any functional limit - persons 18+, all condition</td>
</tr>
<tr>
<td>FLCARRY</td>
<td>Num</td>
<td>3</td>
<td>How difficult lift/carry 10 lbs w/o spec equip</td>
</tr>
<tr>
<td>FLCLIMB</td>
<td>Num</td>
<td>3</td>
<td>How difficult walk up 10 steps w/o spec equip</td>
</tr>
<tr>
<td>FLGRASP</td>
<td>Num</td>
<td>3</td>
<td>How difficult grasp objects w/o spec equip</td>
</tr>
<tr>
<td>FLPUSH</td>
<td>Num</td>
<td>3</td>
<td>How difficult push large object w/o spec equip</td>
</tr>
<tr>
<td>FLREACH</td>
<td>Num</td>
<td>3</td>
<td>How difficult reach over head w/o spec equip</td>
</tr>
<tr>
<td>FLRELAX</td>
<td>Num</td>
<td>3</td>
<td>How difficult do acts to relax w/o spec equip</td>
</tr>
<tr>
<td>FLSHOP</td>
<td>Num</td>
<td>3</td>
<td>How difficult go out to events w/o spec equip</td>
</tr>
<tr>
<td>FLSIT</td>
<td>Num</td>
<td>3</td>
<td>How difficult sit for 2 hrs w/o spec equip</td>
</tr>
<tr>
<td>FLSOCL</td>
<td>Num</td>
<td>3</td>
<td>How difficult social activities w/o spec equip</td>
</tr>
<tr>
<td>FLSTAND</td>
<td>Num</td>
<td>3</td>
<td>How difficult stand for 2 hrs w/o spec equip</td>
</tr>
<tr>
<td>FLSTOOP</td>
<td>Num</td>
<td>3</td>
<td>How difficult stoop/bend/kneel w/o spec equip</td>
</tr>
<tr>
<td>FLWALK</td>
<td>Num</td>
<td>3</td>
<td>How difficult walk 1/4 mile w/o spec equip</td>
</tr>
<tr>
<td>HISCOD_I</td>
<td>Num</td>
<td>3</td>
<td>Combined race/ethnicity recode</td>
</tr>
<tr>
<td>HISCODI2</td>
<td>Num</td>
<td>3</td>
<td>Race/ethnicity recode</td>
</tr>
<tr>
<td>HISCODI3</td>
<td>Num</td>
<td>3</td>
<td>Race/ethnicity recode</td>
</tr>
<tr>
<td>HISPAN_P</td>
<td>Num</td>
<td>3</td>
<td>Hispanic subgroup detail</td>
</tr>
<tr>
<td>HISPCODE</td>
<td>Num</td>
<td>3</td>
<td>Combined race/ethnicity recode</td>
</tr>
<tr>
<td>HISP_CODEC</td>
<td>Num</td>
<td>3</td>
<td>Combined race/ethnicity recode</td>
</tr>
<tr>
<td>HYPEV</td>
<td>Num</td>
<td>3</td>
<td>Ever been told you had hypertension</td>
</tr>
<tr>
<td>MIEV</td>
<td>Num</td>
<td>3</td>
<td>Ever been told you had a heart attack</td>
</tr>
<tr>
<td>OCCUP2</td>
<td>Num</td>
<td>3</td>
<td>Simple occupation classification</td>
</tr>
<tr>
<td>OCCUP2A</td>
<td>Num</td>
<td>3</td>
<td>Simple occupation classification</td>
</tr>
<tr>
<td>OCCUPN2</td>
<td>Num</td>
<td>3</td>
<td>Simple occupation classification</td>
</tr>
<tr>
<td>PHSTAT</td>
<td>Num</td>
<td>3</td>
<td>Reported health status</td>
</tr>
<tr>
<td>Variable</td>
<td>Type</td>
<td>Level</td>
<td>Description</td>
</tr>
<tr>
<td>-----------</td>
<td>-------</td>
<td>-------</td>
<td>-------------</td>
</tr>
<tr>
<td>PLAADL</td>
<td>Num</td>
<td>3</td>
<td>Does -- need help w/ADL?</td>
</tr>
<tr>
<td>PLAIADL</td>
<td>Num</td>
<td>3</td>
<td>Does - - need help w/chores, shop, etc.</td>
</tr>
<tr>
<td>PLAWKLIM</td>
<td>Num</td>
<td>3</td>
<td>Is - - limited kind/amt of work (health)</td>
</tr>
<tr>
<td>PLAWKNOW</td>
<td>Num</td>
<td>3</td>
<td>Is - - unable to work due to health prob</td>
</tr>
<tr>
<td>PLIMANY</td>
<td>Num</td>
<td>3</td>
<td>Is -- limited in ANY (other) WAY?</td>
</tr>
<tr>
<td>PSU</td>
<td>Num</td>
<td>3</td>
<td>PSU for variance estimation</td>
</tr>
<tr>
<td>RACEREC</td>
<td>Num</td>
<td>3</td>
<td>Race Recode</td>
</tr>
<tr>
<td>SEX</td>
<td>Num</td>
<td>3</td>
<td>Sex</td>
</tr>
<tr>
<td>SMKSTAT2</td>
<td>Num</td>
<td>3</td>
<td>Smoking Status: Recode 2</td>
</tr>
<tr>
<td>SRVY_YR</td>
<td>Num</td>
<td>4</td>
<td>Survey Year</td>
</tr>
<tr>
<td>STRATUM</td>
<td>Num</td>
<td>4</td>
<td>Stratum for variance estimation</td>
</tr>
<tr>
<td>STREV</td>
<td>Num</td>
<td>3</td>
<td>Ever been told you had a stroke</td>
</tr>
<tr>
<td>WTFA_SA</td>
<td>Num</td>
<td>8</td>
<td>Weight - Final Annual</td>
</tr>
</tbody>
</table>
## II. MEPS

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Type</th>
<th>Length</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADSMOK42</td>
<td>Num</td>
<td>8</td>
<td>SAQ: CURRENTLY SMOKE</td>
</tr>
<tr>
<td>AGE</td>
<td>Num</td>
<td>8</td>
<td>AGE</td>
</tr>
<tr>
<td>ANGIDX</td>
<td>Num</td>
<td>8</td>
<td>ANGINA DIAGNOSIS (&gt;17)</td>
</tr>
<tr>
<td>ANGIDX53</td>
<td>Num</td>
<td>8</td>
<td>ANGINA DIAGNOSIS (&gt;17) - RD 5/3</td>
</tr>
<tr>
<td>ARTHDX</td>
<td>Num</td>
<td>8</td>
<td>ARTHRITIS DIAGNOSIS (&gt;17)</td>
</tr>
<tr>
<td>ARTHDX53</td>
<td>Num</td>
<td>8</td>
<td>ARTHRITIS DIAGNOSIS (&gt;17) - RD 5/3</td>
</tr>
<tr>
<td>ASTMHDX</td>
<td>Num</td>
<td>8</td>
<td>ASTHMA DIAGNOSIS</td>
</tr>
<tr>
<td>ASTHDX53</td>
<td>Num</td>
<td>8</td>
<td>ASTHMA DIAGNOSIS - RD 5/3</td>
</tr>
<tr>
<td>CHDDX</td>
<td>Num</td>
<td>8</td>
<td>CORONARY HRT DISEASE DIAG (&gt;17)</td>
</tr>
<tr>
<td>CHDDX53</td>
<td>Num</td>
<td>8</td>
<td>CORONARY HRT DISEASE DIAG (&gt;17) - RD 5/3</td>
</tr>
<tr>
<td>CIND31</td>
<td>Num</td>
<td>8</td>
<td>CONDENSED INDUSTRY CODE RD 3/1 CMJ</td>
</tr>
<tr>
<td>CIND42</td>
<td>Num</td>
<td>8</td>
<td>CONDENSED INDUSTRY CODE RD 4/2 CMJ</td>
</tr>
<tr>
<td>COCCP31</td>
<td>Num</td>
<td>8</td>
<td>CONDENSED OCCUPATION CODE RD 3/1 CMJ</td>
</tr>
<tr>
<td>COCCP42</td>
<td>Num</td>
<td>8</td>
<td>CONDENSED OCCUPATION CODE RD 4/2 CMJ</td>
</tr>
<tr>
<td>DIABDX</td>
<td>Num</td>
<td>8</td>
<td>DIABETES DIAGNOSIS (&gt;17)</td>
</tr>
<tr>
<td>DIABDX53</td>
<td>Num</td>
<td>8</td>
<td>DIABETES DIAGNOSIS - RD 5/3</td>
</tr>
<tr>
<td>DUPERSID</td>
<td>Char</td>
<td>8</td>
<td>PERSON ID (DUID + PID)</td>
</tr>
<tr>
<td>EDUCYEAR</td>
<td>Num</td>
<td>8</td>
<td>COMPLETED YEARS OF EDUCATION</td>
</tr>
<tr>
<td>EDUCYR</td>
<td>Num</td>
<td>8</td>
<td>YEARS OF EDUC WHEN FIRST ENTERED MEPS</td>
</tr>
<tr>
<td>EMPHDX</td>
<td>Num</td>
<td>8</td>
<td>EMPHYSEMA DIAGNOSIS (&gt;17)</td>
</tr>
<tr>
<td>EMPHDX53</td>
<td>Num</td>
<td>8</td>
<td>EMPHYSEMA DIAGNOSIS (&gt;17) - RD 5/3</td>
</tr>
<tr>
<td>EMPST13</td>
<td>Num</td>
<td>8</td>
<td>EMPLOYMENT STATUS</td>
</tr>
<tr>
<td>EMPST42</td>
<td>Num</td>
<td>8</td>
<td>EMPLOYMENT STATUS RD 4/2</td>
</tr>
<tr>
<td>EQU42</td>
<td>Num</td>
<td>8</td>
<td>SAQ: EQ-5D PREFERENCE BASED INDEX</td>
</tr>
<tr>
<td>ERTOT</td>
<td>Num</td>
<td>8</td>
<td># EMERGENCY ROOM VISITS 96</td>
</tr>
<tr>
<td>EVRETIRE</td>
<td>Num</td>
<td>8</td>
<td>PERSON HAS EVER RETIRED</td>
</tr>
<tr>
<td>EVRWRK</td>
<td>Num</td>
<td>8</td>
<td>EVER WORKED FOR PAY AS OF December</td>
</tr>
<tr>
<td>HIBPDX</td>
<td>Num</td>
<td>8</td>
<td>HIGH BLOOD PRESSURE DIAG (&gt;17)</td>
</tr>
<tr>
<td>HIBPDX53</td>
<td>Num</td>
<td>8</td>
<td>HIGH BLOOD PRESSURE DIAG (&gt;17) - RD 5/3</td>
</tr>
<tr>
<td>INSCOV</td>
<td>Num</td>
<td>8</td>
<td>HEALTH INSURANCE COVERAGE INDICATOR</td>
</tr>
<tr>
<td>IPDIS</td>
<td>Num</td>
<td>8</td>
<td># HOSPITAL DISCHARGES</td>
</tr>
<tr>
<td>IPNGTD</td>
<td>Num</td>
<td>8</td>
<td># NIGHTS IN HOSP FOR DISCHARGES</td>
</tr>
<tr>
<td>IPZERO</td>
<td>Num</td>
<td>8</td>
<td># ZERO-NIGHT HOSPITAL STAYS</td>
</tr>
<tr>
<td>MCS42</td>
<td>Num</td>
<td>8</td>
<td>SAQ:MENTAL COMPONET SUMMRY SF12 IMPUTED</td>
</tr>
<tr>
<td>MIDX</td>
<td>Num</td>
<td>8</td>
<td>HEART ATTACK (MI) DIAG (&gt;17)</td>
</tr>
<tr>
<td>Variable</td>
<td>Type</td>
<td>Value</td>
<td>Description</td>
</tr>
<tr>
<td>------------</td>
<td>-------</td>
<td>-------</td>
<td>-------------------------------------------------------</td>
</tr>
<tr>
<td>MIDX53</td>
<td>Num</td>
<td>8</td>
<td>HEART ATTACK (MI) DIAG (&gt;17) - RD 5/3</td>
</tr>
<tr>
<td>OBTOTV</td>
<td>Num</td>
<td>8</td>
<td># OFFICE BASED PROVIDER VISITS</td>
</tr>
<tr>
<td>OCCCAT31</td>
<td>Num</td>
<td>8</td>
<td>OCCUPATION GROUP RD 3/1 CMJ</td>
</tr>
<tr>
<td>OCCCAT42</td>
<td>Num</td>
<td>8</td>
<td>OCCUPATION GROUP RD 4/2 CMJ</td>
</tr>
<tr>
<td>OHRTDX</td>
<td>Num</td>
<td>8</td>
<td>OTHER HEART DISEASE DIAG (&gt;17)</td>
</tr>
<tr>
<td>OHRTDX53</td>
<td>Num</td>
<td>8</td>
<td>OTHER HEART DISEASE DIAG (&gt;17) - RD 5/3</td>
</tr>
<tr>
<td>OPTOTV</td>
<td>Num</td>
<td>8</td>
<td># OUTPATIENT PROVIDER VISITS</td>
</tr>
<tr>
<td>PCS42</td>
<td>Num</td>
<td>8</td>
<td>SAQ: PHYSICAL COMPNT SUMMARY SF12 IMPUTED</td>
</tr>
<tr>
<td>POVCAT</td>
<td>Num</td>
<td>8</td>
<td>POVERTY STATUS</td>
</tr>
<tr>
<td>PSU9609</td>
<td>Num</td>
<td>8</td>
<td>COMBINED VARIANCE PSU: 1996-2009</td>
</tr>
<tr>
<td>QOL_WT</td>
<td>Num</td>
<td>8</td>
<td>SAQ FILE WEIGHT</td>
</tr>
<tr>
<td>RACETHNX</td>
<td>Num</td>
<td>8</td>
<td>RACE/ETHNICITY (EDITED/IMPUTED)</td>
</tr>
<tr>
<td>RACEX</td>
<td>Num</td>
<td>8</td>
<td>RACE (EDIT/IMPUTED)</td>
</tr>
<tr>
<td>SEX</td>
<td>Num</td>
<td>8</td>
<td>SEX (EDITED/IMPUTED)</td>
</tr>
<tr>
<td>STRA9609</td>
<td>Num</td>
<td>8</td>
<td>COMBINED VARIANCE STRATUM: 1996-2009</td>
</tr>
<tr>
<td>STRKDX</td>
<td>Num</td>
<td>8</td>
<td>STROKE DIAGNOSIS (&gt;17)</td>
</tr>
<tr>
<td>STRKDX53</td>
<td>Num</td>
<td>8</td>
<td>STROKE DIAGNOSIS (&gt;17) - RD 5/3</td>
</tr>
<tr>
<td>TOTEXP</td>
<td>Num</td>
<td>8</td>
<td>TOTAL HEALTH CARE EXP</td>
</tr>
<tr>
<td>YEAR</td>
<td>Char</td>
<td>4</td>
<td>SURVEY YEAR</td>
</tr>
</tbody>
</table>
Appendix B: Source Code

SAS: Specific Aim 1

libname NHIS 'C:\datasets';
options nofmterr;
data NHIS.paper1;

/* age domains */
if 18 <= age <= 25 then age_g65=1; else if (26 <= age <= 64) then age_g65=2; else if (age >= 65) then age_g65=3;

/* multimorbidity - "cond2" variable */
if HYPEV=1 then HTN=1; else if HYPEV=2 then HTN=0; else HTN=.;
if CHDEV=1 then CHD=1; else if CHDEV=2 then CHD=0; else CHD=.;
if ANGEV=1 then angina=1; else if ANGEV=2 then angina=0; else angina=.;
if MIEV=1 then MI=1; else if MIEV=2 then MI=0; else MI=.;
    if (CHD=1 or angina=1 or MI=1) then heart=1; else if (CHD=. and angina=. and MI=.) then heart=.; else heart=0;
if STREV=1 then stroke=1; else if STREV=2 then stroke=0; else stroke=.;
if EPHEV=1 then emphys=1; else if EPHEV=2 then emphys=0; else emphys=.;
if AASMEV=1 then asthma=1; else if AASMEV=2 then asthma=0; else asthma=.;
if CANEV=1 then cancer=1; else if CANEV=2 then cancer=0; else cancer=.;
if DIBEV=1 then diabetes=1; else if DIBEV=2 then diabetes=0; else diabetes=.;
cond=sum(HTN, heart, stroke, emphys, asthma, cancer, diabetes);
cond2=(cond>=2);

/* self-rated health */
if (1 <= PHSTAT < 4) then hlthstat=1; /* good-excellent */
else if (4 <= PHSTAT < 6) then hlthstat=2; /* poor-fair */

/* smoking status */
if SMKSTAT2 in (1 2) then smoke=2; /* 2 = current, 1 = former, 0 = never;*/
    else if SMKSTAT2=3 then smoke=1;
    else if SMKSTAT2=4 then smoke=0;
    else if SMKSTAT2=5 then do;
        if smoke_cur in (1 2) then smoke=2;
        else if smoke_cur=0 then smoke=1; end;
    else smoke=.;

/* alcohol consumption */
if 1997 <= srvy_yr <= 2003 then do;
    If ALCSTAT1=1 then alcohol=0; /* 0 = lifetime abstainer, 1 = former drinker,*/
        2 = current drinker, no heavy drinking, 3 = current drinker and history of heavy drinking in the last year;
else if ALCSTAT1=2 then alcohol=1;
else if (ALCSTAT1=3) then do;
    if 1<=ALC5UPYR<=365 then alcohol=3;
else alcohol=2; end;
else alcohol=; end;
else if 2004<=srvy_yr then do;
    If ALCSTAT=1 then alcohol=0;
else if ALCSTAT in (2 3 4) then alcohol=1;
else if ALCSTAT in (5 6 7 8 9) then do;
    if 1<=ALC5UPYR<=365 then alcohol=3;
else alcohol=2; end;
else alcohol=.
end;
/*functional limitations */
If FLWALK in (1 2 3 4) then walk=1; else if FLWALK=0 then walk=0; else if FLWALK in (6 7 8 9) then walk=.;
If FLCLIMB in (1 2 3 4) then climb=1; else if FLCLIMB=0 then climb=0; else if FLCLIMB in (6 7 8 9) then climb=.;
If FLSTAND in (1 2 3 4) then stand=1; else if FLSTAND=0 then stand=0; else if FLSTAND in (6 7 8 9) then stand=.;
If FLSIT in (1 2 3 4) then sit=1; else if FLSIT=0 then sit=0; else if FLSIT in (6 7 8 9) then sit=.;
If FLSTOOP in (1 2 3 4) then stoop=1; else if FLSTOOP=0 then stoop=0; else if FLSTOOP in (6 7 8 9) then stoop=.;
If FLREACH in (1 2 3 4) then reach=1; else if FLREACH=0 then reach=0; else if FLREACH in (6 7 8 9) then reach=.;
If FLGRASP in (1 2 3 4) then grasp=1; else if FLGRASP=0 then grasp=0; else if FLGRASP in (6 7 8 9) then grasp=.;
If FLCARRY in (1 2 3 4) then carry=1; else if FLCARRY=0 then carry=0; else if FLCARRY in (6 7 8 9) then carry=.;
If FLPUSH in (1 2 3 4) then push=1; else if FLPUSH=0 then push=0; else if FLPUSH in (6 7 8 9) then push=.;
If FLSHOP in (1 2 3 4) then shop=1; else if FLSHOP=0 then shop=0; else if FLSHOP in (6 7 8 9) then shop=.;
If FLSOCL in (1 2 3 4) then soc=1; else if FLSOCL=0 then soc=0; else if FLSOCL in (6 7 8 9) then soc=.;
If FLRELAX in (1 2 3 4) then relax=1; else if FLRELAX=0 then relax=0; else if FLRELAX in (6 7 8 9) then relax=.;
If FLA1AR=1 then ADL=1; else if FLA1AR=2 then ADL=0; else if FLA1AR=3 then ADL=.;
limit=sum(walk, climb, stand, sit, stoop, reach, grasp, carry, push, shop, soc, relax);
if limit in (0 1) then limit2=0; else if limit>=2 then limit2=1;

**utility index according to Livingston and Ko, 2002;**
if plaadl = 1 then plaadl_ = 0; *1=needs help with personal care;
if plaiadl = 1 then plaiadl_ = .2; *1=needs help with routine needs;
if plawknow = 1 then plawknow_ = .4; *1=unable to work due to health problem;
if plawklim = 1 then plawklim_ = .65; *0,1=limited kind/amount of work;
if plimany = 1 then plimany_ = .75; *0,1=limited in any (other) way?

if n(plaadl_, plaiadl_, plawknow_, plawklim_, plimany_) >= 1
    then sas = min(plaadl_, plaiadl_, plawknow_, plawklim_, plimany_);
else if plaadl^=1 and plaiadl^=1 and plawknow^=1 and plawklim^=1 and plimany^=1
    then sas = 1;

if PHSTAT = 1 then phs = 1.00;
else if PHSTAT = 2 then phs = .85;
else if PHSTAT = 3 then phs = .70;
else if PHSTAT = 4 then phs = .30;
else if PHSTAT = 5 then phs = 0;

if sas^=, and phs ^=, then do;
m=(.41*phs) + (.41*sas) + (.18*phs*sas);
UI = .10 + (.9*m);
end;

if ui > .80 then UI_cat=5;
else if ui > .60 then UI_cat=4;
else if ui > .40 then UI_cat=3;
else if ui > .20 then UI_cat=2;
else if ui > .00 then UI_cat=1;

if 0<=UI<=0.481940 then ui20=0; else if ui>0.481940 then ui20=1; else ui20=;
  * ui20=0 - below 20th percentile for HALex, ui20=1 - above 20th percentile;

/* education */
if educ<=12 then education=1; /* 1<HS, 2=HS, 3>HS */
else if (educ=13 or educ=14) then education=2;
else if 15<=educ<=21 then education=3;

/* race/ethnicity */
if srvy_yr=1997 then do;
  if 0<=hispan_p<=11 then raceth=1; /* raceth=1: hispanic 2 = White non-hisp; 3 = Black; 4 = Other */
  else raceth=racerec+1;
end;
else if srvy_yr=1998 then raceth=hispcod;
else if srvy_yr=1999 then raceth=hispcodr;
else if srvy_yr in(2000 2001 2002) then raceth=hiscodi_i;
else if srvy_yr in(2003 2004 2005) then raceth=hiscodi2;
else if srvy_yr in(2006 2007 2008 2009 2010 2011 2012) then do;
raceth=hiscodi3;
if raceth=5 then raceth=4;
end;

/* employment */
if 1997<=srvy_yr<=2000 then do;
    if doinglw in (1 2) then employed=1; else if doinglw in (3 4) then employed=0;
else employed=.; end;
else if 2001<=srvy_yr then do;
    if doinglw in (1 2) then employed=1; else if doinglw in (3 4 5) then employed=0;
else employed=.; end;

/* occupation */
if (1997<=srvy_yr<=2003) then do;
    industry41=indstry1;
    occupation13=occup2; end;
else if (srvy_yr=2004) then do;
    industry41=indstr1a;
    occupation13=occup2a; end;
if (2004<=srvy_yr) then do;
    industry23=indstrn2;
    occupation23=occupn2; end;

if 1997<=srvy_yr<=2004 then do;
    if occupation13 in (1 2 3 4 5) then krieger=1; *white collar;
    else if occupation13 in (6 7 8) then krieger=2; *services;
    else if occupation13=9
    then krieger=3; *farm;
    else if occupation13 in (10 11 12 13) then krieger=4; end; *blue collar;

else if srvy_yr>=2005 then do;
    if occupation23 in (1,2,3,4,5,7,8,9,10,16,17) then krieger=1;
    else if occupation23 in (6,11,12,13,14,15) then krieger=2;
    else if occupation23 in (18) then krieger=3;
    else if occupation23 in (19,20,21,22) then krieger=4;
    end; else krieger=.;

if employed=0 then krieger5=0; else krieger5=krieger;

keep employed sector krieger smoke alcohol insurance education raceth hlthstat age_g65
age psu strat wt sex
krieger5 cond cond2 HTN heart stroke emphys asthma cancer diabetes chd angina MI
srvy_yr
limit limit2 ADL UI UI_cat ui20 ;
run;

proc contents data=NHIS.paper1 position ;
run;

/*Variables in Creation Order
# Variable Type Len  Label
1 SRVY_YR Num 4  Survey Year
2 SEX Num 3  Sex
3 PSU Num 3  PSU for variance estimation
4 wt Num 8  sample adult file weight/15
5 strat Num 8  stratum for variance estimation
6 age Num 8  age
7 employed Num 8  employment status
8 sector Num 8  industry sector
9 krieger Num 8  krieger occupation category
10 smoke Num 8  smoking status
11 alcohol Num 8  alcohol consumption
12 education Num 8  education level
13 raceth Num 8  race/ethnicity
14 hlthstat Num 8  self-rated health
15 age_g65 Num 8  age group
16 HTN Num 8  hypertension lifetime diagnosis
17 CHD Num 8  coronary heart disease lifetime diagnosis
18 angina Num 8  angina lifetime diagnosis
19 MI Num 8  myocardial infarction lifetime diagnosis
20 heart Num 8  heart disease lifetime diagnosis
21 stroke Num 8  stroke lifetime diagnosis
22 emphys Num 8  emphysema lifetime diagnosis
23 asthma Num 8  asthma lifetime diagnosis
24 cancer Num 8  cancer lifetime diagnosis
25 diabetes Num 8  diabetes lifetime diagnosis
26 cond Num 8  number of chronic conditions
27 cond2 Num 8  multiple chronic conditions/multimorbidity
28 ADL Num 8  Any functional limitation, all conditions
29 limit Num 8  number of functional limitations
30 limit2 Num 8  multiple functional limitations present
31 UI Num 8  HALex score
32 UI_cat Num 8  HALex quintile
33 ui20 Num 8  in HALex 20th percentile
34 krieger5 Num 8  5-category employment/occupation classification */

Proc surveylogist data=NHIS.paper1;
  cluster psu;
  strata strat;
  weight wt;
  domain age_g65;
  class raceth (param=ref ref='2') education (param=ref ref='1') sex (param=ref ref='2')
            krieger5 (param=ref ref='0') smoke (param=ref ref='0') alcohol (param=ref ref='0');
model ui20=education raceth sex age smoke alcohol krieger5 sryv_yr sex*krieger5/ rsq;
run;

Proc surveylogist data=NHIS.paper1;
  cluster psu;
  strata strat;
  weight wt;
  domain age_g65;
  class raceth (param=ref ref='2') education (param=ref ref='1') sex (param=ref ref='2')
  krieger5 (param=ref ref='0') smoke (param=ref ref='0') alcohol (param=ref ref='0');
  model hlthstat (descending)=education raceth sex age smoke alcohol krieger5 sryv_yr
  sex*krieger5/rsq;
run;

Proc surveylogist data=NHIS.paper1;
  cluster psu;
  strata strat;
  weight wt;
  domain age_g65;
  class raceth (param=ref ref='2') education (param=ref ref='1') sex (param=ref ref='2')
  krieger5 (param=ref ref='0') smoke (param=ref ref='0') alcohol (param=ref ref='0');
  model cond2 (descending)=education raceth sex age smoke alcohol krieger5 sryv_yr
  sex*krieger5/rsq;
run;

Proc surveylogist data=NHIS.paper1;
  cluster psu;
  strata strat;
  weight wt;
  domain age_g65;
  class raceth (param=ref ref='2') education (param=ref ref='1') sex (param=ref ref='2')
  krieger5 (param=ref ref='0') smoke (param=ref ref='0') alcohol (param=ref ref='0');
  model limit2 (descending)=education raceth sex age smoke alcohol krieger5 sryv_yr
  sex*krieger5/ rsq;
run;

Proc surveylogist data=NHIS.paper1;
  cluster psu;
  strata strat;
  weight wt;
  domain age_g65*sex;
  class raceth (param=ref ref='2') education (param=ref ref='1') sex (param=ref ref='2')
  krieger5 (param=ref ref='0') smoke (param=ref ref='0') alcohol (param=ref ref='0');
  model cond2 (descending)=education raceth sex age smoke alcohol krieger5 sryv_yr /rsq;
run;
**Proc surveylogist** data=NHIS.paper1;
   cluster psu;
   strata strat;
   weight wt;
   domain age_g65;
   class raceth (param=ref ref='2') education (param=ref ref='1') sex (param=ref ref='2')
          krieger5 (param=ref ref='0') smoke (param=ref ref='0') alcohol (param=ref ref='0');
   model ui20=education raceth sex age smoke alcohol krieger5 srvy_yr / rsq ;
   run;

**Proc surveylogist** data=NHIS.paper1;
   cluster psu;
   strata strat;
   weight wt;
   domain age_g65;
   class raceth (param=ref ref='2') education (param=ref ref='1') sex (param=ref ref='2')
          krieger5 (param=ref ref='0') smoke (param=ref ref='0') alcohol (param=ref ref='0');
   model hlthstat (descending)=education raceth sex age smoke alcohol krieger5 srvy_yr /
                       rsq ;
   run;

**Proc surveylogist** data=NHIS.paper1;
   cluster psu;
   strata strat;
   weight wt;
   domain age_g65;
   class raceth (param=ref ref='2') education (param=ref ref='1') sex (param=ref ref='2')
          krieger5 (param=ref ref='0') smoke (param=ref ref='0') alcohol (param=ref ref='0');
   model limit2 (descending)=education raceth sex age smoke alcohol krieger5 srvy_yr /
                       rsq ;
   run;
SAS: code used to create MEPS dataset (Specific Aims 2 & 3)

libname MEPS 'C:\files';
libname datasets 'C:\datasets';
options nofmterr;

Data datasets.MEPS;
set MEPS.pooled_cons;

ind9031=CIND31; ind9042=CIND42; ind9053=CIND53;
ind0231=INDCAT31; ind0242=INDCAT42; ind0253=INDCAT53;
occ9031=COCCP31; occ9042=COCCP42; occ9053=COCCP53;
occ0231=OCCCAT31; occ0242=OCCCAT42; occ0253=OCCCAT53;

if EMPST42 in (1 2 3) then employed42=1; else if EMPST42=4 then employed42=0; else employed42=.;
if EMPST13 in (1 2 3) then employed31=1; else if EMPST13=4 then employed31=0; else employed31=.;

if occ9042=-2 then occ9042=occ9031; if occ9053=-2 then occ9053=occ9042;
* "-2" means occupational information collected in the previous round. ;
if occ0242=-2 then occ0242=occ0231; if occ0253=-2 then occ0253=occ0242;
if 1996<=year<=2001 then do;
/* '31' occupations are from the beginning of the year are for use in expenditures/utilization analyses */
if employed31=0 then krieger31=0; *unemployed;
else if employed31=1 then do;
if occ9031 in (1 2 3 4) then krieger31=1;
*1=white collar;
else if occ9031 in (5 6 7 9) then krieger31=2;
*2=blue collar;
else if occ9031 in (10 11) then krieger31=3;
*3=farmer;
else if (occ9031=8 and employed31=1) then krieger31=4; end; *4=service;
/* '42 occupations were measured at the same time as the quality of life - for use in QOL analyses */
if employed42=0 then krieger42=0; *unemployed;
else if employed42=1 then do;
if occ9042 in (1 2 3 4) then krieger42=1;
*1=white collar;
else if occ9042 in (5 6 7 9) then krieger42=2;
*2=blue collar;
else if occ9042 in (10 11) then krieger42=3; *3=farmer;
else if occ9042=8 then krieger42=4; *4=service;
end; end;
else if year>=2002 then do;
  if employed31=0 then krieger31=0;
  else if employed31=1 then do;
    if occ0231 in (1 2 4 5) then krieger31=1;
    else if occ0231 in (7 8) then krieger31=2;
    else if occ0231 in (6) then krieger31=3;
    else if occ0231=3 then krieger31=4; end;
  if employed42=0 then krieger42=0;
  else if employed42=1 then do;
    if occ0242 in (1 2 4 5) then krieger42=1;
    else if occ0242 in (7 8) then krieger42=2;
    else if occ0242 in (6) then krieger42=3;
    else if occ0242=3 then krieger42=4; end;
end;

  if age in (-3 -1) then age=.;
if age>85 then age=85; * up to year 2000 top-coded at 90, starting in 2001 only at 85.
I top-coded everything at 85.;

if RACETHNX=1 then raceth=3; *hispanic for all years;
else do;
  if 1996<=year<=2001 then do;
    if RACEX=5 then raceth=1; *white;
    else if RACEX=4 then raceth=2; *black;
    else if RACEX in (1 2 3) then raceth=4; *other;
    end;
    else if year>=2002 then do;
      if RACEX=1 then raceth=1;
      else if RACEX=2 then raceth=2;
      else if RACEX in (3 4 5 6) then raceth=4;
    end;
end;

if 1999<=year<=2004 then educ_year=EDUCYEAR; else if year>=2005 then educ_year=EDUCYR;
if educ_year<0 then educ_year=.;

if 2000<=year<=2006 then do;
  if ASTHDX53=1 then asthma=1; else if ASTHDX53=2 then asthma=0; else asthma=.;
  if HIBPDX53=1 then HTN=1; else if HIBPDX53=2 then HTN=0; else HTN=.;
if (CHDDX53=1 or MIDX53=1 or OHRTDX53=1 or ANGIDX53=1) then heart=1;
else if (CHDDX53=2 and MIDX53=2 and OHRTDX53=2 and ANGIDX53=2) then heart=0; else heart=.
if DIABDX53=1 then diabetes=1; else if DIABDX53=2 then diabetes=0; else diabetes=.
if EMPHDX53=1 then emphysema=1; else if EMPHDX53=2 then emphysema=0; else emphysema=.
if STRKDX53=1 then stroke=1; else if STRKDX53=2 then stroke=0; else stroke=.
if ARTHDX53=1 then arthrit=1; else if ARTHDX53=2 then arthrit=0; else arthrit=.
end;
if 2007<=year<=2009 then do;
if ASTHDX=1 then asthma=1; else if ASTHDX=2 then asthma=0; else asthma=.
if ARTHDX=1 then arthrit=1; else if ARTHDX=2 then arthrit=0; else arthrit=.
if DIABDX=1 then diabetes=1; else if DIABDX=2 then diabetes=0; else diabetes=.
if EMPHDX=1 then emphysema=1; else if EMPHDX=2 then emphysema=0; else emphysema=.
if STRKDX=1 then stroke=1; else if STRKDX=2 then stroke=0; else stroke=.
if HIBPDX=1 then HTN=1; else if HIBPDX=2 then HTN=0; else HTN=.
if (CHDDX=1 or MIDX=1 or OHRTDX=1 or ANGIDX=1) then heart=1;
else if (CHDDX=2 and MIDX=2 and OHRTDX=2 and ANGIDX=2) then heart=0; else heart=.
end;
if ADSMOK42=1 then smoke=1; else if ADSMOK42=2 then smoke=0; else smoke=.
if age>=65 then domain=1; else domain=0;
chronic=sum(asthma, HTN, heart, diabetes, emphysema, stroke); * 2000-2009 - arthritis not included;
if 2001<=year<=2009 then chronic_7_01=sum(asthma, HTN, heart, diabetes, emphysema, stroke, arthrit); * only 2001-2009;
if EVRETIRE<1 then EVRETIRE=.; *1 YES, 2 NO;
if PCS42 in (-9 -1) then PCS42=.;
if MCS42 in (-9 -1) then MCS42=.;
if EQU42 in (-9 -1) then EQU42=.;

if krieger31=0 then do; kr0=1; kr1=0; kr2=0; kr3=0; kr4=0;
else if krieger31=1 then do; kr0=0; kr1=1; kr2=0;
  kr3=0; kr4=0; end;
else if krieger31=2 then do; kr0=0; kr1=0; kr2=1;
  kr3=0; kr4=0; end;
else if krieger31=3 then do; kr0=0; kr1=0; kr2=0;
  kr3=1; kr4=0; end;
else if krieger31=4 then do; kr0=0; kr1=0; kr2=0;
  kr3=0; kr4=1; end;

if krieger42=0 then do; krq0=1; krq1=0; krq2=0; krq3=0;
  krq4=0; end;
else if krieger42=1 then do; krq0=0; krq1=1; krq2=0;
  krq3=0; krq4=0; end;
else if krieger42=2 then do; krq0=0; krq1=0; krq2=1;
  krq3=0; krq4=0; end;
else if krieger42=3 then do; krq0=0; krq1=0; krq2=0;
  krq3=1; krq4=0; end;
else if krieger42=4 then do; krq0=0; krq1=0; krq2=0;
  krq3=0; krq4=1; end;

if inscov=1 then do; ins_priv=1; ins_pub=0; ins_none=0;
else if inscov=2 then do; ins_priv=0; ins_pub=1;
  ins_none=0; end;
else if inscov=3 then do; ins_priv=0; ins_pub=0;
  ins_none=1; end;

prov_vis=sum (OBTOTV, OPTOTV); * total number of provider visits - outpatient hospital + office based;

if raceth=1 then WNH=1 ; else WNH=0; *white;
if raceth=2 then BNH=1; else BNH=0; *black;
if raceth=3 then Hisp=1; else Hisp=0; *hispanic;
if raceth=4 then reth0=1; else reth0=0; *other;

if QOL_WT^=. ; *SAQ weight ;

drop EDUCYEAR [long list of variables] OBDRV ; run;

proc contents data=datasets.MEPS; run;
/*Alphabetic List of Variables and Attributes

<table>
<thead>
<tr>
<th>#</th>
<th>Variable</th>
<th>Type</th>
<th>Len</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>49</td>
<td>BNH</td>
<td>Num</td>
<td>8</td>
<td>BLACK NON-HISPANIC DUMMY VARIABLE</td>
</tr>
<tr>
<td>1</td>
<td>DUPERSID</td>
<td>Char</td>
<td>8</td>
<td>PERSON ID (DUID + PID)</td>
</tr>
<tr>
<td>15</td>
<td>EMPST42</td>
<td>Num</td>
<td>8</td>
<td>EMPLOYMENT STATUS RD 4/2</td>
</tr>
<tr>
<td>19</td>
<td>EQU42</td>
<td>Num</td>
<td>8</td>
<td>SAQ: EQ-5D PREFERENCE BASED INDEX</td>
</tr>
<tr>
<td>10</td>
<td>ERTOT</td>
<td>Num</td>
<td>8</td>
<td># EMERGENCY ROOM VISITS 96</td>
</tr>
<tr>
<td>16</td>
<td>EVRETIRED</td>
<td>Num</td>
<td>8</td>
<td>PERSON HAS EVER RETIRED</td>
</tr>
<tr>
<td>7</td>
<td>EVRWRK</td>
<td>Num</td>
<td>8</td>
<td>EVER WORKED FOR PAY AS OF December</td>
</tr>
<tr>
<td>50</td>
<td>Hisp</td>
<td>Num</td>
<td>8</td>
<td>HISPANIC DUMMY VARIABLE</td>
</tr>
<tr>
<td>8</td>
<td>INSCOV</td>
<td>Num</td>
<td>8</td>
<td>HEALTH INSURANCE COVERAGE</td>
</tr>
<tr>
<td>12</td>
<td>IPDIS</td>
<td>Num</td>
<td>8</td>
<td># HOSPITAL DISCHARGES</td>
</tr>
<tr>
<td>13</td>
<td>IPNGTD</td>
<td>Num</td>
<td>8</td>
<td># NIGHTS IN HOSP FOR DISCHARGES</td>
</tr>
<tr>
<td>11</td>
<td>IPZERO</td>
<td>Num</td>
<td>8</td>
<td># ZERO-NIGHT HOSPITAL STAYS</td>
</tr>
<tr>
<td>18</td>
<td>MCS42</td>
<td>Num</td>
<td>8</td>
<td>SAQ:MENTAL COMPONENT SUMMARY SF12</td>
</tr>
<tr>
<td>22</td>
<td>OBTOTV</td>
<td>Num</td>
<td>8</td>
<td># OFFICE BASED PROVIDER VISITS</td>
</tr>
<tr>
<td>23</td>
<td>OPTOTV</td>
<td>Num</td>
<td>8</td>
<td># OUTPATIENT PROVIDER VISITS</td>
</tr>
<tr>
<td>17</td>
<td>PCS42</td>
<td>Num</td>
<td>8</td>
<td>SAQ:PHYSICAL COMPONENT SUMMARY SF12</td>
</tr>
<tr>
<td>6</td>
<td>POVCAT</td>
<td>Num</td>
<td>8</td>
<td>POVERTY STATUS</td>
</tr>
<tr>
<td>3</td>
<td>PSU9609</td>
<td>Num</td>
<td>8</td>
<td>COMBINED VARIANCE PSU: 1996-2009</td>
</tr>
<tr>
<td>20</td>
<td>QOL_WT</td>
<td>Num</td>
<td>8</td>
<td>SAQ FILE WEIGHT</td>
</tr>
<tr>
<td>4</td>
<td>SEX</td>
<td>Num</td>
<td>8</td>
<td>SEX (EDITED/IMPUTED)</td>
</tr>
<tr>
<td>2</td>
<td>STRA9609</td>
<td>Num</td>
<td>8</td>
<td>COMBINED VARIANCE STRATUM: 1996-2009</td>
</tr>
<tr>
<td>9</td>
<td>TOTEXP</td>
<td>Num</td>
<td>8</td>
<td>TOTAL HEALTH CARE EXP</td>
</tr>
<tr>
<td>48</td>
<td>WNH</td>
<td>Num</td>
<td>8</td>
<td>WHITE NON-HISPANIC DUMMY VARIABLE</td>
</tr>
<tr>
<td>26</td>
<td>age</td>
<td>Num</td>
<td>8</td>
<td>AGE</td>
</tr>
<tr>
<td>29</td>
<td>arthrit</td>
<td>Num</td>
<td>8</td>
<td>ARTHRITIS DIAGNOSIS</td>
</tr>
<tr>
<td>32</td>
<td>chronic</td>
<td>Num</td>
<td>8</td>
<td># CHRONIC CONDITIONS, ARTHRITIS</td>
</tr>
<tr>
<td>33</td>
<td>chronic_7_01</td>
<td>Num</td>
<td>8</td>
<td># CHRONIC CONDITIONS, ARTHRITIS INCLUDED, 2000-2009</td>
</tr>
<tr>
<td>31</td>
<td>domain</td>
<td>Num</td>
<td>8</td>
<td>DOMAIN</td>
</tr>
<tr>
<td>28</td>
<td>educ_year</td>
<td>Num</td>
<td>8</td>
<td>NUMBER OF YEARS OF EDUCATION</td>
</tr>
<tr>
<td>24</td>
<td>employed31</td>
<td>Num</td>
<td>8</td>
<td>EMPLOYMENT STATUS IN ROUND 1 OR 3</td>
</tr>
<tr>
<td>46</td>
<td>ins_none</td>
<td>Num</td>
<td>8</td>
<td>INSURANCE STATUS DUMMY VAR: NONE</td>
</tr>
<tr>
<td>44</td>
<td>ins_priv</td>
<td>Num</td>
<td>8</td>
<td>INSURANCE STATUS DUMMY VAR: PUBLIC</td>
</tr>
<tr>
<td>45</td>
<td>ins_pub</td>
<td>Num</td>
<td>8</td>
<td>INSURANCE STATUS DUMMY VAR: PUBLIC</td>
</tr>
<tr>
<td>34</td>
<td>kr0</td>
<td>Num</td>
<td>8</td>
<td>OCCUPATION DUMMY ROUND 3 OR 1: UNEMPLOYED</td>
</tr>
<tr>
<td>35</td>
<td>kr1</td>
<td>Num</td>
<td>8</td>
<td>OCCUPATION DUMMY ROUND 3 OR 1: WHITE COLLAR</td>
</tr>
</tbody>
</table>
data out;
set datasets.MEPS; format _all_;
file "C:\MEPS_exp_diss.dat" delimiter = ",";
put DUPERSID STRA9609 PSU9609 SEX POVCAT INSCOV year EVRETIRE TOTEXP ERTOT IPDIS IPNGTD OBTOTV OPTOTV prov_vis ins_priv ins_pub ins_none QOL_WT age raceth educ_year chronic_7_01 domain smoke kr0 kr1 kr2 kr3 kr4 WNH BNH Hisp rethO ;
run;

data out;
set datasets.MEPS; format _all_;
file "C:\MEPS_QOL_diss.dat" delimiter = ",";
put DUPSERSID STRA9609 PSU9609 SEX POVCAT INSCOV year EVRETIRE PCS42 MCS42 EQU42 QOL_WT age_dec krieger42 raceth educ_year chronic domain smoke krq0 krq1 krq2 krq3 krq4 ins_priv ins_pub ins_none chronic_7_01 wt_equ domainE WNH BNH Hisp rethO ;
run;
**Mplus: Specific Aim 2**

Note: models testing mediation paths are not presented in this appendix. The only difference between those and the ones presented was that in the mediation models employment/occupation and health insurance variables were not declared categorical.

**TITLE:** EuroQual-5D employed vs unemployed

**DATA:**
```
FILE IS "D:\MEPS_QOL_chron.dat";
```

**VARIABLE:**
```
NAMES ARE DUPERSID STRA9609 PSU9609 SEX POVCAT INSCOV year EVRETIRe
PCS42 MCS42 EQU42 QOL_WT age_dec krieger4 raceth educ_year chronic_old
domain smoke kr0 kr1 kr2 kr3 kr4 ins_priv ins_pub ins_none
chronic wt_equ domainE WNH BNH Hisp rethO;
MISSING are .
```

**USEVARIABLES ARE**
```
SEX POVCAT ins_priv ins_none EQU42 age_dec BNH Hisp rethO educ_year chronic smoke kr0
```

**ANALYSIS:** TYPE Complex;
ESTIMATOR is MLR;
Integration= MONTECARLO (100);

MODEL:  EQU42 on SEX  POVCAT ins_priv  ins_none kr0 chronic age_dec  
educ_year smoke;
kr0 on SEX  age_dec  educ_year chronic smoke;
chronic on SEX  age_dec  BNH Hisp rethO  
educ_year smoke;
POVCAT on SEX  chronic age_dec  BNH Hisp rethO  
educ_year kr0 smoke;
ins_priv-ins_none on SEX  POVCAT chronic age_dec  BNH Hisp rethO  
educ_year smoke;

OUTPUT:  SAMPSTAT standardized residual cinterval svalues;
TITLE: EuroQual-5D all occupations vs white collar

DATA:
FILE IS "D:\MEPS_QOL_chron.dat";

VARIABLE:
NAMES ARE DUPERSID STRA9609 PSU9609 SEX POVCAT INSCOV year EVRETIRE
PCS42 MCS42 EQU42 QOL_WT age_dec krieger4 raceth educ_year chronic_old
domain smoke kr0 kr1 kr2 kr3 kr4 ins_priv ins_pub ins_none
chronic wt_equ domainE WNH BNH Hisp rethO;
MISSING are .;

USEVARIABLES ARE SEX POVCAT ins_priv ins_none EQU42
age_dec BNH Hisp rethO educ_year chronic smoke kr0 kr2 kr3 kr4;
categorical are povcat ins_priv ins_none kr0 kr2 kr3 kr4
BNH Hisp rethO sex smoke;

stratification is STRA9609;
cluster is PSU9609;
weight is wt_equ;

subpopulation is domainE == 1;

ANALYSIS: TYPE Complex;
ESTIMATOR is MLR;
Integration= MONTECARLO (100);

MODEL: EQU42 on SEX POVCAT ins_priv ins_none kr0 - kr4 chronic age_dec
educ_year smoke;
kr0 - kr4 on SEX  age_dec  BNH Hisp rethO  
educ_year chronic smoke;
chronic on SEX  age_dec  BNH Hisp rethO  
educ_year smoke;
POVCAT on SEX  chronic age_dec  BNH Hisp rethO  
educ_year kr0 - kr4 smoke;
ins_priv - ins_none on SEX  POVCAT chronic age_dec  BNH Hisp rethO  
educ_year smoke;

OUTPUT: SAMPSTAT standardized residual cinterval svalues;
TITLE: MCS Employed vs Unemployed

DATA:
FILE IS "D:\MEPS_QOL_chron.dat";

VARIABLE:
NAMES ARE DUPERSID STRA9609 PSU9609 SEX POVCAT INSCOV year EVRETIRE
PCS42 MCS42 EQU42 QOL_WT age_dec krieger4 raceth educ_year chronic_old
domain smoke kr0 kr1 kr2 kr3 kr4 ins_priv ins_pub ins_none
chronic wt_equ domainE WNH BNH Hisp rethO;
MISSING are .;
USEVARIABLES ARE SEX POVCAT ins_priv ins_none MCS42
age_dec BNH Hisp rethO educ_year chronic smoke kr0 ;
categorical are povcat ins_priv ins_none kr0 ;
stratification is STRA9609;
cluster is PSU9609;
weight is QOL_WT;
subpopulation is domain == 1;

ANALYSIS: TYPE Complex ;
ESTIMATOR is MLR;
Integration= MONTECARLO (100);
MCONVERGENCE = 1.0E-06
MODEL:    MCS42 on SEX  POVCAT ins_PRIV  ins_NONE kr0 chronic age_dec educ_year BNH Hisp rethO smoke;
kr0 on SEX  age_dec  BNH Hisp rethO educ_year chronic ;
chronic on SEX  age_dec  BNH Hisp rethO educ_year smoke;
POVCAT on SEX  chronic age_dec  BNH Hisp rethO educ_year kr0 smoke;
ins_PRIV-ins_NONE on SEX  POVCAT chronic age_dec  BNH Hisp rethO educ_year kr0 ;

OUTPUT:  SAMPSTAT standardized residual  cinterval svalues;
TITLE: MCS all occupations vs white collar

DATA: FILE IS "D:\MEPS_QOL_chron.dat";

VARIABLE:
  NAMES ARE DUPERSID STRA9609 PSU9609 SEX POVCAT INSCOV year EVRETI RE
  PCS42 MCS42 EQU42 QOL_WT age_dec krieger4 raceth educ_year chronic_old
domain smoke kr0 kr1 kr2 kr3 kr4 ins_priv ins_pub ins_none
chronic wt_equ domainE WNH BNH Hisp rethO;

MISSING are .;

USEVARIABLES ARE SEX POVCAT ins_priv ins_none MCS42
  age_dec BNH Hisp rethO educ_year chronic smoke kr0 kr2 kr3 kr4;
  categorical are povcat ins_priv ins_none kr0 kr2 kr3 kr4
     BNH Hisp rethO sex smoke;

  stratification is STRA9609;
  cluster is PSU9609;
  weight is QOL_WT;

  subpopulation is domain == 1;

ANALYSIS: TYPE Complex;
  ESTIMATOR is MLR;
  Integration= MONTECARLO (100);
  MCONVERGENCE = 1.0E-06

MODEL: MCS42 on SEX POVCAT ins_priv ins_none kr0 - kr4 chronic age_dec
educ_year BNH Hisp rethO smoke;
kr0 - kr4 on SEX age_dec BNH Hisp rethO
    educ_year chronic smoke;
chronic on SEX age_dec BNH Hisp rethO
    educ_year smoke;
POVCAT on SEX chronic age_dec BNH Hisp rethO
    educ_year kr0 - kr4 smoke;
ins_priv on SEX POVCAT chronic age_dec BNH Hisp rethO
    educ_year kr0 - kr4 ;
ins_none on SEX POVCAT chronic age_dec BNH Hisp rethO
    educ_year kr0 - kr4 ;

OUTPUT: SAMPSTAT standardized residual cinterval svalues;
TITLE: PCS Employed vs Unemployed

DATA:
   FILE IS "D:\MEPS_QOL_chron.dat";

VARIABLE:
   NAMES ARE DUPERSID STRA9609 PSU9609 SEX POVCAT INSCOV year EVRETIRE
   PCS42 MCS42 EQU42 QOL_WT age_dec krieger4 raceth educ_year chronic_old
domain smoke kr0 kr1 kr2 kr3 kr4 ins_priv ins_pub ins_none
   chronic wt_equ domainE WNH BNH Hisp rethO;
   MISSING are .;

   USEVARIABLES ARE SEX POVCAT ins_priv ins_none PCS42 age_dec BNH Hisp rethO educ_year chronic smoke kr0 ;
   categorical are povcat ins_priv ins_none kr0 ;

   stratification is STRA9609;
   cluster is PSU9609;
   weight is QOL_WT;

   subpopulation is domain == 1;

ANALYSIS: TYPE Complex ;
   ESTIMATOR is MLR;
   Integration= MONTECARLO (100);
   MODEL: PCS42 on SEX POVCAT ins_priv ins_none kr0 chronic age_dec educ_year smoke;
   kr0 on SEX age_dec BNH Hisp rethO
educc_year chronic;
chronic on SEX age_dec BNH Hisp rethO
educc_year smoke;
POVCAT on SEX chronic age_dec BNH Hisp rethO
educc_year kr0 smoke;
ins_priv-ins_none on SEX POVCAT chronic age_dec BNH Hisp rethO
educc_year kr0;

OUTPUT: SAMPSTAT standardized residual CINTERVAL SVALUES TECH10;

TITLE: PCS all occupations vs white collar

DATA:
FILE IS "D:\MEPS_QOL_chron.dat";

VARIABLE:
NAMES ARE DUPERSID STRA9609 PSU9609 SEX POVCAT INSCOV year EVRETIRE
PCS42 MCS42 EQU42 QOL_WT age_dec krieger4 raceth educ_year chronic_old
domain smoke kr0 kr1 kr2 kr3 kr4 ins_priv ins_pub ins_none
chronic wt_equ domainE WNH BNH Hisp rethO;
MISSING are .;

USEVARIABLES ARE SEX POVCAT ins_priv ins_none PCS42 age_dec BNH Hisp rethO educ_year chronic smoke kr0 kr2 kr3 kr4;
categorical are povcat ins_priv ins_none kr0 kr2 kr3 kr4
BNH Hisp rethO sex smoke;

stratification is STRA9609;
cluster is PSU9609;
weight is QOL_WT;

subpopulation is domain == 1;

ANALYSIS: TYPE Complex;
ESTIMATOR is MLR;
Integration= MONTECARLO (100);

MODEL: PCS42 on SEX POVCAT ins_priv ins_none kr0 - kr4 chronic age_dec
educ_year smoke;
kr0 - kr4 on SEX age_dec BNH Hisp rethO
  educ_year chronic smoke;
chronic on SEX age_dec BNH Hisp rethO
  educ_year smoke;
POVCAT on SEX chronic age_dec BNH Hisp rethO
  educ_year kr0 - kr4 smoke;
ins_priv on SEX POVCAT chronic age_dec BNH Hisp rethO
  educ_year kr0 - kr4 ;
ins_none on SEX POVCAT chronic age_dec BNH Hisp rethO
  educ_year kr0 - kr4 ;

OUTPUT: SAMPSTAT standardized residual CINTERVAL SVALUES TECH10 ;
Mplus: Specific Aim 3

TITLE: # nights of hospital stay

DATA:
FILE IS "D:\MEPS_exp_chron.dat";

VARIABLE:
NAMES ARE DUPERSID STRA9609 PSU9609 SEX POVCAT INSCOV year EVRETIRE

   TOTEXP ERTOT IPDIS IPNGTD OBTOTV OPTOTV prov_vis ins_priv ins_pub ins_none QOL_WT age_dec raceth educ_year chronic
domain smoke kr0 kr1 kr2 kr3 kr4 WNH BNH Hisp rethO;
MISSING are .;

USEVARIABLES ARE SEX POVCAT ins_priv ins_none IPNGTD age_dec BNH Hisp rethO educ_year chronic smoke kr0 kr2 kr3 kr4 ;
categorical are povcat insPriv insNone kr0 kr2 kr3 kr4 ;
count is IPNGTD (nbi);

stratification is STRA9609;
cluster is PSU9609;
weight is QOL_WT;

subpopulation is domain == 1;

ANALYSIS: TYPE Complex;
Integration = MONTECARLO (100);
MODEL: IPNGTD on SEX POVCAT kr0 - kr4 chronic
    educ_year BNH Hisp rethO ;
IPNGTD#1 on ins_priv ins_none chronic age_dec
    BNH Hisp rethO ;
kr0 - kr4 on SEX age_dec BNH Hisp rethO
    educ_year chronic smoke;
chronic on SEX age_dec BNH Hisp rethO
    educ_year smoke;
POVCAT on SEX chronic age_dec BNH Hisp rethO
    educ_year kr0 - kr4 smoke;

OUTPUT: SAMPSTAT standardized residual cinterval svalues;
TITLE: # ER visits
DATA:
  FILE IS "D:\MEPS_exp_chron.dat";

VARIABLE:
  NAMES ARE DUPERSID STRA9609 PSU9609 SEX POVCAT INSCOV year EVRETIRE
  TOTEXP ERTOT IPDIS IPNGTD OPTOTV OPTOTV prov_vis ins_priv ins_pub
  ins_none QOL_WT age_dec raceth educ_year chronic
domain smoke kr0 kr1 kr2 kr3 kr4 WNH BNH Hisp rethO;

MISSING are .;

USEVARIABLES ARE SEX POVCAT ins_priv ins_none ERTOT
  age_dec BNH Hisp rethO educ_year chronic smoke kr0 kr2 kr3 kr4 ;

categorical are povcat ins_priv ins_none kr0 kr2 kr3 kr4 ;
count is ERTOT (i);

stratification is STRA9609;
can be PSU9609;
weight is QOL_WT;

subpopulation is domain == 1;

ANALYSIS: TYPE Complex ;
  Integration= MONTECARLO (100);

MODEL: ERTOT on kr0 - kr4 chronic ;
  ERTOT#1 on POVCAT ins_priv ins_none chronic age_dec;
  kr0 - kr4 on SEX age_dec BNH Hisp rethO
educ_year chronic smoke;
chronic on SEX age_dec BNH Hisp rethO
educ_year smoke;

OUTPUT: SAMPSTAT standardized residual cinterval svalues;
TITLE: Total Expenditures

DATA:
  FILE IS "C:\MEPS_exp_chron.dat";

VARIABLE:
  NAMES ARE DUPERSID STRA9609 PSU9609 SEX POVCAT INSCOV year EVRETIRED TOTEXP ERTOT IPDIS IPNGTD OBTOTV OPTOTV prov_vis ins_priv ins_pub ins_none QOL_WT age_dec raceth educ_year chronic domain smoke kr0 kr1 kr2 kr3 kr4 WNH BNH Hisp rethO;
  MISSING are .;

  USEVARIABLES ARE SEX POVCAT ins_priv ins_none TOTEXP age_dec BNH Hisp rethO educ_year chronic smoke kr0 kr2 kr3 kr4 ;

  categorical are povcat ins_priv ins_none kr0 kr2 kr3 kr4 ;
  count is TOTEXP (nbi);

stratification is STRA9609;
cluster is PSU9609;
weight is QOL_WT;

subpopulation is domain == 1;

ANALYSIS: TYPE Complex ;
  Integration= MONTECARLO (100);

MODEL: TOTEXP on ins_priv ins_none kr0 - kr4 chronic age_dec educ_year BNH Hisp rethO smoke;
       TOTEXP#1 on SEX POVCAT ins_priv ins_none*1.670 chronic age_dec
educ_year B Hannah Hispanic smoke;
kr0 - kr4 on SEX age_dec B Hannah Hispanic smoke;
educ_year chronic smoke;
chronic on SEX age_dec B Hannah Hispanic smoke;
educ_year smoke;
POVCAT on SEX chronic age_dec B Hannah Hispanic smoke;
educ_year kr0 - kr4 smoke;
ins_priv on SEX POVCAT age_dec B Hannah Hispanic chronic
educ_year kr0 - kr4;
ins_none on SEX POVCAT age_dec B Hannah Hispanic chronic
educ_year kr0 - kr4;

OUTPUT: SAMPSTAT standardized residual cinterval svalues;