Economie d’avant garde

Dissertation No. 3

2006

ON THE CAUSE OF INCREASED LONGEVITY AND URBAN SPRAWL:
A MACROECONOMIC APPROACH

by

Richard M. H. Suen
On the Cause of Increased Longevity and Urban Sprawl:

A Macroeconomic Approach

by

Ming Hon Suen

Submitted in Partial Fulfillment
of the
Requirements for the Degree
Doctor of Philosophy

Supervised by
Professor Jeremy Greenwood

Department of Economics
The College
Arts & Sciences

University of Rochester
Rochester, New York
2006
Curriculum Vitae

The author was born in Hong Kong on August 9th, 1977. He attended The Chinese University of Hong Kong from 1996 to 1999, and graduated with a Bachelor of Social Science degree in 1999. He came to the University of Rochester in the Fall of 2001 and began graduate studies in Economics. He pursued his research in Macroeconomics under the direction of Professor Jeremy Greenwood and received the Master of Arts degree from the University of Rochester in 2005.
Acknowledgements

I owe a great debt to my advisor, Jeremy Greenwood, who gave me continuous support and encouragement during the past five years. His superb guidance is indispensable for the completion of this work. I am also grateful to Mark Bils for his helpful comments and suggestions. I would like to thank my family for giving me tremendous support throughout all these years.
Abstract

Chapter 1  Medical Spending and Longevity

The second half of the twentieth century recorded a rapid growth in health care spending and a significant increase in life expectancy. This paper hypothesizes that the combination of technological progress in medical treatment and rising incomes is the driving force behind these two trends. Using a stochastic, multi-period overlapping-generations model as the analytical vehicle, this paper argues that the rapid growth in medical spending is not driven by factors associated with market structures or insurance opportunities, but by factors underlying the production and accumulation of health. According to this model, improvements in medical treatment and rising incomes can explain all of the increase in medical spending and more than 60% of the increase in life expectancy at age 25 during the second half of the twentieth century.

Chapter 2  The Welfare Implications of Increased Medical Spending

The rapid growth in medical spending over the period 1950-2001 was accompanied by a remarkable expansion in insurance coverage. Earlier studies contended
that due to incentive problems in health insurance, the increase in medical spending represents a significant welfare loss. The objective of this chapter is to re-examine this welfare issue. A complete-market version of the economy presented in Chapter 1 is used as a benchmark for welfare analysis. This is then compared to the incomplete-market economy presented in the first chapter. An equivalent variation measure is used to quantify the extent of welfare loss due to market incompleteness.

Chapter 3 Suburbanization and the Automobile

During the period 1910-1970, an increasing fraction of urban population in the U.S. chose to live outside central cities. This was also a time when automobile was introduced and widely adopted. The objective of this paper is to assess quantitatively the relationship between the two. To achieve this, a simple model is constructed in which people can choose where to live and whether or not to buy a car. The calibrated version of the model is able to explain about 70% of average car-ownership during 1910-1970 and the suburbanization trend.
# Table of Contents

Curriculum Vitae ................................................................. ii  
Acknowledgements .............................................................. iii 
Abstract ................................................................................. iv  
Table of Contents ................................................................. vi  
List of Figures .......................................................................... ix 
List of Tables ............................................................................ xi 

Foreword .................................................................................... 1

1 Medical Spending and Longevity ............................................ 2  
1.1 Introduction ........................................................................... 2  
1.2 History .................................................................................. 5  
  1.2.1 Some Facts ........................................................................ 5  
  1.2.2 Historical Discussions .................................................... 8  
1.3 The Model ............................................................................. 13  
  1.3.1 The Environment ............................................................ 13  
  1.3.2 The Benchmark Economy in 1950 ................................. 15  
  1.3.3 The Benchmark Economy in 2001 .................................. 21  
  1.3.4 The Economy with No Insurance ................................. 24  
1.4 Quantitative Analysis .......................................................... 25  
1.5 Findings ............................................................................... 30  
  1.5.1 Benchmark Economy ..................................................... 30  
  1.5.2 Economy with No Insurance .......................................... 33  
  1.5.3 Robustness ..................................................................... 35
2 The Welfare Implications of Increased Medical Spending

3 Suburbanization and the Automobile
3.4  Quantitative Analysis  
   3.4.1  Parameters to set  
   3.4.2  Estimation  

3.5  Findings  
   3.5.1  Baseline Results  
   3.5.2  Discussions  
   3.5.3  Counterfactual Experiments  

3.6  Concluding Remarks  

3.7  Appendix  

3.8  Figures and Tables  

Bibliography
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.1</td>
<td>Real Per-capita Personal Medical Expenditures, 1930-2001.</td>
<td>42</td>
</tr>
<tr>
<td>Figure 1.2</td>
<td>Share of Medical Spending in GDP, 1930-2001.</td>
<td>42</td>
</tr>
<tr>
<td>Figure 1.3</td>
<td>Components of Personal Medical Expenditures, 1930-2001.</td>
<td>43</td>
</tr>
<tr>
<td>Figure 1.4</td>
<td>Life Expectancy at Birth and Age-adjusted Mortality Rate, 1900-2001.</td>
<td>43</td>
</tr>
<tr>
<td>Figure 1.5</td>
<td>Survival Probability at Age 25 and above, 1949-51 and 2001.</td>
<td>44</td>
</tr>
<tr>
<td>Figure 1.6</td>
<td>Distribution of Personal Medical Expenditures by Source of Funds, 1950-2001.</td>
<td>44</td>
</tr>
<tr>
<td>Figure 1.7</td>
<td>Percentage Distribution of Deaths by Age, 1900-2001.</td>
<td>45</td>
</tr>
<tr>
<td>Figure 1.8</td>
<td>Age-adjusted Mortality Rates (per 100,000 population) of Diseases of Heart and Cancer, 1900-2001.</td>
<td>45</td>
</tr>
<tr>
<td>Figure 1.9</td>
<td>Average Annual Incidence Rate of Coronary Heart Disease, 1987-2000.</td>
<td>46</td>
</tr>
<tr>
<td>Figure 1.10</td>
<td>Average Annual Incidence Rate of Cancers, 1989-2001.</td>
<td>46</td>
</tr>
<tr>
<td>Figure 1.11</td>
<td>Life Expectancy by Age, 1950.</td>
<td>47</td>
</tr>
<tr>
<td>Figure 1.12</td>
<td>Life Expectancy by Age, 2001.</td>
<td>47</td>
</tr>
<tr>
<td>Figure 1.13</td>
<td>Survival Probabilities in 1950.</td>
<td>48</td>
</tr>
<tr>
<td>Figure 1.14</td>
<td>Survival Probabilities in 2001.</td>
<td>48</td>
</tr>
<tr>
<td>Figure</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>------------</td>
<td>-----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Figure 1.15</td>
<td>Ratio of Medical Spending to Income, 1950.</td>
<td>49</td>
</tr>
<tr>
<td>Figure 1.16</td>
<td>Ratio of Medical Spending to Income, 2001.</td>
<td>49</td>
</tr>
<tr>
<td>Figure 2.1</td>
<td>Relationship between End-of-Period Health and Medical Care.</td>
<td>89</td>
</tr>
<tr>
<td>Figure 3.1</td>
<td>Population Density Distributions for Atlanta in 1910, 1950 and 1970.</td>
<td>131</td>
</tr>
<tr>
<td>Figure 3.2</td>
<td>Population Density Distributions for Chicago in 1910 and 1950.</td>
<td>131</td>
</tr>
<tr>
<td>Figure 3.3</td>
<td>Number of Registered Automobiles per capita, 1900-1993.</td>
<td>132</td>
</tr>
<tr>
<td>Figure 3.4</td>
<td>Percentage of Families who Own a Car by Number of Cars Owned, 1950-1995.</td>
<td>132</td>
</tr>
<tr>
<td>Figure 3.5</td>
<td>Car Ownership by Income Quintile, 1952-1965.</td>
<td>133</td>
</tr>
<tr>
<td>Figure 3.6</td>
<td>Time Cost of New Automobiles (in Logarithm), 1906-2000.</td>
<td>133</td>
</tr>
<tr>
<td>Figure 3.7</td>
<td>Car Ownership in 1962 by Distance from Center of Central City.</td>
<td>134</td>
</tr>
<tr>
<td>Figure 3.8</td>
<td>Car Ownership in 12 Largest SMSA’s by Place of Residence, 1961 and 1971.</td>
<td>134</td>
</tr>
<tr>
<td>Figure 3.9</td>
<td>Highway and Street Stock per Capita, 1925-94.</td>
<td>135</td>
</tr>
<tr>
<td>Figure 3.10</td>
<td>Capital Stock per Capita of Equipment and Structures for Intercity and Local Passenger Transit, 1947-1970.</td>
<td>135</td>
</tr>
<tr>
<td>Figure 3.11</td>
<td>Public Transit Ridership per Capita, 1902-1970.</td>
<td>136</td>
</tr>
</tbody>
</table>
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1.1</td>
<td>Age-specific Death Rate [per 100,000 population for specified group].</td>
<td>50</td>
</tr>
<tr>
<td>Table 1.2</td>
<td>Death Rates [per 100,000 population] for Specific Infectious Diseases, 1900-1980.</td>
<td>50</td>
</tr>
<tr>
<td>Table 1.3</td>
<td>Baseline Parameterization</td>
<td>51-52</td>
</tr>
<tr>
<td>Table 1.4</td>
<td>Selected Real-World Statistics.</td>
<td>53</td>
</tr>
<tr>
<td>Table 1.5</td>
<td>Main Findings from the Benchmark Model.</td>
<td>54</td>
</tr>
<tr>
<td>Table 1.6</td>
<td>Changes in Average Life Expectancy between 1950 and 2001.</td>
<td>55</td>
</tr>
<tr>
<td>Table 1.7</td>
<td>Ratio of Medical Spending to Income by Age.</td>
<td>55</td>
</tr>
<tr>
<td>Table 1.8</td>
<td>Benchmark Economy vs Economy with No Insurance.</td>
<td>56</td>
</tr>
<tr>
<td>Table 1.9</td>
<td>Effects of Changing the Coefficient of Relative Risk Aversion.</td>
<td>56</td>
</tr>
<tr>
<td>Table 1.10</td>
<td>Effects of Changing the Price of Medical Care in 1950 ($p_{1950}$).</td>
<td>57</td>
</tr>
<tr>
<td>Table A1</td>
<td>Real Per-Person Medical Spending by Age, 1950-2000.</td>
<td>58</td>
</tr>
<tr>
<td>Table A2</td>
<td>Real Per-Person Medical Spending for Selected Age Groups, 1950 and 2000.</td>
<td>59</td>
</tr>
<tr>
<td>Table A3</td>
<td>Age-specific Death Rates [per 100,000 population for each group], 1950.</td>
<td>60</td>
</tr>
<tr>
<td>Table A4</td>
<td>Age-specific Death Rates [per 100,000 population for each group], 2001.</td>
<td>61</td>
</tr>
<tr>
<td>Table A5</td>
<td>The Cost-sharing Structure of Medicare, 1966-2001.</td>
<td>62</td>
</tr>
<tr>
<td>Table</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>-----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Table 2.1</td>
<td>Findings: Complete vs Incomplete Markets.</td>
<td>90</td>
</tr>
<tr>
<td>Table 3.1</td>
<td>Coefficients for the Negative Exponential Function, $d = ae^{-bx}$, for Chicago, 1900-1950.</td>
<td>137</td>
</tr>
<tr>
<td>Table 3.2</td>
<td>Coefficients for the Negative Exponential Function, $d = ae^{-bx}$, for Atlanta, 1900-1970.</td>
<td>137</td>
</tr>
<tr>
<td>Table 3.3</td>
<td>Mean Gradients for Forty-One U.S. Cities that were Metropolitan Districts in 1900.</td>
<td>138</td>
</tr>
<tr>
<td>Table 3.4</td>
<td>Number of Registered Vehicles per Person Aged 20-64, 1900-1970.</td>
<td>138</td>
</tr>
<tr>
<td>Table 3.5</td>
<td>Net Capital Expenditures of Urban Transit Properties, 1890-1950.</td>
<td>139</td>
</tr>
<tr>
<td>Table 3.6</td>
<td>Baseline Parameterization</td>
<td>140</td>
</tr>
<tr>
<td>Table 3.7</td>
<td>Baseline Results</td>
<td>141</td>
</tr>
<tr>
<td>Table 3.8</td>
<td>Counterfactual Experiments</td>
<td>142</td>
</tr>
</tbody>
</table>
Foreword

The original title of Chapter 1 is “Technological Advances and the Growth in Health Care Spending.” The third chapter of this dissertation is a joint work with Karen A. Kopecky.
Chapter 1

Medical Spending and Longevity

1.1 Introduction

It is well documented that over the latter half of the twentieth century, Americans faced a rapid growth in health care spending and a significant increase in longevity. Over the period 1950-2001, real per-capita personal medical spending increased by a factor of 9.7, while life expectancy at age 25 extended by 6.8 years. Despite the wide disparity in health care systems, similar patterns were observed in other OECD countries over the same time period. This suggests that at least part of these changes was driven by factors that are common across borders. The same time period also witnessed an influx of new techniques and technologies that revolutionized the practice of medicine. Procedures that once existed only in fantasy, like open-heart surgery, organ transplants, in vitro fertilization, etc., are now in widespread practice. This chapter hypothesizes that the combination of technological progress in medical treatment and rising incomes is the driving force behind the increase in both medical spending and longevity.

To address the question in hand, a stochastic, multi-period overlapping generations model is adopted as the analytical vehicle. The core of the analysis is the
individual’s demand for health. Health is formulated as a stock that prolongs life, while medical spending is considered as an investment into this stock. At each age, a person faces two types of uncertainty. First, there is a given probability of getting sick. Diseases are considered to be negative shocks that lower a person’s health status. The severeness of illness can be alleviated by transforming medical care into new units of health. Second, at the end of each age, the person will face a certain probability of dying. This probability is endogenously determined by the health status. Health is demanded because it raises the survival probability in all states and all future periods. Using medical care as input, health is produced via a health production function. Medical care refers to commodities such as a hospital stay or a physician visit. The current state of the art of medical technology is embodied in the medical treatment received during the hospital stay or physician visit. Medical treatment refers to a set of diagnostic and therapeutic procedures targeted on certain diseases or conditions. In the model, this is captured by the form of the production function. Medical technology is improved when innovations in medical treatment raise the marginal product of medical care in producing health. In the current framework, such an innovation is expenditure raising for two reasons. First, more effective treatment encourages use at all ages. This is particularly prominent among the elderly with poor health. When productivity is low, receiving medical treatment does not bring much improvement in their chances of survival and so medical spending tends to be low (or even zero). The advent of new, effective technologies brings new hopes and medical spending increases significantly as a result. Second, when lives are extended by the new technologies, additional spending will be incurred during the added years of life. In other words, total medical expenditures increase because people live longer.

The same topic is studied in the recent work by Hall and Jones (2004). There
are two major differences between this and the current work. First, Hall and Jones believe that income growth by itself can explain the expansion in medical spending. The key assumption behind their argument is that consumption in medical services is more income elastic than other form of consumption. The current paper emphasizes both technological improvement and income growth. In the quantitative analysis, it is shown that income growth alone is not enough to generate the observed increase in medical spending. Another major difference is that in Hall and Jones, there is no uncertainty in medical spending and hence no role for health insurance. In the current study, the relationship between medical spending and insurance coverage is examined.

Health insurance is central for the question in hand. Earlier studies contended that the growth in health care spending is mainly due to the fact that the U.S. population is overinsured.\(^1\) This claim is re-examined in this paper. This is done by comparing two model economies with different degrees of insurance opportunity. In the benchmark model, two sources of health insurance are available via which consumers can protect themselves against unexpected need for medical expenses. First, there is a private insurance market in which reimbursement health insurance contracts are traded. Second, a public health insurance program, similar to the Medicare program, is provided by the government. In the comparison, this benchmark model is used to represent the real world. In the second model economy, both public and private health insurance are removed. The two models are identical in all other aspects. The difference between the two thus capture the impact of expanded insurance coverage on medical spending.

In the quantitative analysis, two steady states are constructed, based on the benchmark model, to represent the U.S. economy in 1950 and 2001. The cal-

\(^1\)Their argument is detailed in Chapter 2.
ibrated model is able to match the observed changes in medical spending and insurance coverage. It also yields reasonable predictions on the life-cycle behavior of consumption and medical spending. When the same set of parameters is imposed on the economy with no insurance, the following predictions can be obtained. First, a large increase in medical spending occurs over time even when consumers are cut off from all insurance markets. Second, technological progress in medical treatment, coupled with rising incomes, can explain all the increase in medical spending, and more than 60% of the increase in life expectancy at age 25 during the second half of the twentieth century. These findings illustrate the main idea of this paper: The increase in medical spending is not driven by factors associated with insurance opportunities or market structures, but by factors underlying the production and accumulation of health.

The remainder of this chapter is structured as follows. The next section reviews the trends in medical spending, life expectancy and insurance coverage during the second half of the twentieth century. It also briefly describes several major changes that the U.S. health care sector has gone through over the past hundred years. Section 1.3 presents the model economies. Section 1.4 discusses the calibration procedure. The main findings are reported in section 1.5. This is followed by some concluding remarks in section 1.6.

1.2 History

1.2.1 Some Facts

Figure 1.1 illustrates the dramatic increase in real per-capita personal medical expenditures during the twentieth century. In 1950 a typical American spent

---

2 Personal medical expenditures include spending on hospital services; physician, dentist and other professional services; prescription drugs and medical equipment; nursing home services,
$448 (in constant 2001 dollars) on medical care. This increased by a factor of 9.7 over the next fifty years and reached $4,361 in 2001.\footnote{In the quantitative analysis, only those aged 25 or above will be considered. Thus it is of interest to consider the average spending of this group. Since these statistics are not readily obtainable, they are computed using the data reported in the technical appendix of Meara, White and Cutler (2004). Details of the computation can be found in the Appendix. For 1950 and 2001, the average medical spending of those aged 25 years or above were $584 (in constant 2001 dollars) and $5,140, respectively. This implies an average annual growth rate of 4.36\%.} During the same period, real per-capita GDP increased only by a factor of 2.9. The result is a rising share of medical spending in GDP as shown in Figure 1.2. Over the period 1950-2001 the share of medical spending in GDP increased from 3.7\% to 12.2\%. From the figure it is obvious that the rising trend began at some point around 1950. This is one of the reasons why this paper focuses on the latter half of the twentieth century.\footnote{Another reason is that formal medicine played a significant role in reducing mortality during the same time period but not earlier on. Readers are referred to the historical discussions for further details.} When Medicare and Medicaid were enacted in 1966 the share of medical spending had already increased to 4.9\%. Another way to assess the rising importance of health care spending is to consider its share in personal consumption expenditures.\footnote{Data on personal consumption expenditures (PCE) are obtained from the Bureau of Economic Analysis, National Income and Product Accounts. The medical care component of PCE is replaced by the personal medical expenditures described in footnote 2. This is because the former includes net costs of health insurance.} In 1950 medical spending accounted for a mere 5.6\% of total personal consumption expenditures, much smaller than the shares on food (27.5\%) and housing (11.1\%). As depicted in Figure 1.3, this share remained almost constant during the early decades and started gaining momentum only in the 1950s. The share of medical spending had already reached 7.9\% by the year 1966. In 2001 the share of medical care in total consumption expenditures etc. These exclude expenses on the following items: net costs of health insurance, medical facilities construction, program administration, government public health activities and research. For 1930-1960, the data source is Worthington (1975). For 1960-2001, the data are obtained from the U.S. Centers for Medicare and Medicaid Services, “Health Accounts”, \url{<http://cms.hhs.gov/statistics/nhe/default.asp>}. These data are then divided by the total civilian population and deflated by the GDP deflator.
was 17.3%, exceeding the share on either food (13.5%) or housing (15.0%). Over the same time period, medical spending has increased for all age groups but at different paces. The increase was particularly prominent among the elderly (those aged 65 or above). Over the period 1950-2000, real per-person spending among the elderly grew at an average annual rate of 5.4%, whereas the corresponding growth rate for the non-elderly was 3.8%.\footnote{See Appendix for details.}

The tremendous increase in medical expenditures was accompanied by (i) a decline in mortality and (ii) an expansion in insurance coverage. The declining trend in age-adjusted mortality and the corresponding increase in life expectancy are shown in Figure 1.4.\footnote{Mortality rates are \textit{age-adjusted} when the underlying age composition is held constant as in certain base year. This makes comparisons across years meaningful. For the data reported here, the base year is 2000. Unless otherwise specified, all mortality rates reported in this work are age-adjusted and come from the same source: Centers for Disease Control and Prevention, National Center for Health Statistics, \texttt{<http://www.cdc.org/nchs/about/major/dvs/mortdata.htm>}. Data on life expectancy can be found in \textit{National Vital Statistics Reports}, v.52, no.14. Data on life expectancy are constructed using cross-sectional data on age-specific death rates for the designated years. These data include deaths due to events that are not directly related to health, such as accident, homicide and suicide. Although these three causes accounted for only 8.3% of all deaths in 1950 and 6.3% in 2001, they accounted for a large fraction of deaths among teenagers and young adults. Readers are referred to Tables A3 and Table A4 for detailed statistics. For the purpose of this chapter, it is desirable to consider deaths caused by health-related events only. When deaths due to accident, homicide and suicide are removed, life expectancy at age 25 was 47.7 years in 1950 and 54.2 years in 2001.\footnote{Deaths due to accident, homicide and suicide are removed when preparing these figures.}}

In the middle of the twentieth century, mortality rate stood at 1,446 per 100,000 population. Over the next 50 years, mortality dropped by 40% and became 855 per 100,000 population. An average American at age 25 in 1950 could expect to live 46.6 more years. This increased to 53.4 years by 2001.\footnote{Deaths due to accident, homicide and suicide are removed when preparing these figures.} An alternative way to describe the increase in longevity is to consider the proportion of population that survives to a certain age. Figure 1.5 compares the probability of being alive in various ages (conditional on being alive at age 25) between 1949-51 and 2001.\footnote{Deaths due to accident, homicide and suicide are removed when preparing these figures.}
Insurance coverage also expanded markedly during the post-war era. Between 1950 and 2001 the proportion of total civilian population covered by health insurance increased from 58% to 85.7%. At the same time, private insurance became an increasingly important source of payment for personal health care [see Figure 1.6]. In 1950 68.3% of total personal health care expenses were paid directly by the consumers, whereas only 8.5% were paid by private insurance. Half a century later, out-of-pocket expenditures accounted for only 17% of the total spending, while 35% were paid by private insurance. Since the introduction of Medicare and Medicaid, the U.S. government has assumed a considerably larger role in financing the provision of health care. In 2001 43.4% of total medical spending were paid by the government.

### 1.2.2 Historical Discussions

#### Before 1950

One of the most remarkable developments in the twentieth century is the rapid decline in mortality. In 1900 mortality rate stood at 2,518 per 100,000 population, while life expectancy at birth was just 47.3 years. This was a time when infectious diseases were rampant. Pneumonia, influenza and tuberculosis topped the list of leading killers at that time. These, together with diphtheria, measles, scarlet fever and whooping cough, accounted for 27% of all deaths. The influenza epidemic of 1918, which pushed the death rate to 2,542 per 100,000 population, demonstrated how devastating these diseases could be if left unchecked. The prevalence of infectious diseases also made young children’s lives more vulnerable. Infant mortality

---


rate was lamentably high at the beginning of the last century. One out of every six babies born could not survive their first year [see Table 1.1]. Children under 5 years of age accounted for 30% of all deaths while persons aged 65 and above accounted for only 24%. Over the ensuing fifty years, mortality rate fell by 42.6% while average lifespan extended by 20.9 years. The decline in mortality during early childhood was particularly impressive. By 1950 less than 10% of all deaths were for people under five years of age. The changing age composition of deaths is illustrated in Figure 1.7.

What caused these changes in mortality? In particular, what was the role played by modern medical science? Effective control over infectious diseases was the driving force behind the reduction in mortality during the first half of the twentieth century. This was primarily the result of rising living standards and improved public health measures, such as proper sewage disposal and purified water.\textsuperscript{12} Formal medicine only played an ancillary role in the battle against infectious diseases at that time.

By the end of the nineteenth century, medical researchers had already identified the causes of various infectious diseases. But little was known on how to treat them. When effective medicine, such as penicillin and other antibiotics, finally emerged in the 1940s, infectious disease mortality was already under control [see Table 1.2]. The quality of medical practice during the earlier decades is also an issue of concern. Regulations on medical practice and medical education were nonexistent before 1920. This provided a breeding-ground for a large number of incompetent practitioners produced by proprietary medical schools. These schools would admit anyone who could afford the tuition fees without regard to

\textsuperscript{12}Interested readers are referred to Dowling (1977) for a detailed and non-technical account on the battle against infectious diseases.
their academic background. Many specialists in the profession were self-named and poorly trained.\textsuperscript{13} Even for well-trained physicians, there was not much they could do in terms of diagnosis and treatment. A physician recalled the state of the art in internal medicine in 1930 as follows:

“We had only a few drugs to prescribe then, and our diagnostic aids were simple and crude. By contrast, a half century later, some thousands of drugs are available for use in the practice of internal medicine, most of which have been proved effective. ... Similarly with diagnostic procedures, there was little to be done up to 1930... We had no chemical tests for blood levels of enzymes, electrolytes, gases, or hormones, and none of our modern imaging techniques were available.” Beeson (1980).

Major reforms in medical education and practice had been undertaken since. Academic prerequisites for medical education were specified. Clinical teaching became part of the medical curriculum and internship was required for graduation. A system of board certification was established to regulate the professional accreditation of specialists. All these have had far-reaching impact on the quality of physicians and the public’s perception towards them.

The first half of the twentieth century also witnessed a rapid development in private health insurance. The first generation of health insurance contracts, which were developed during the 1930s, took the form of prepayment plans organized by hospital associations, medical professions and community-consumer groups. Suffering from the accumulation of unpaid medical bills during the Great Depression, hospitals and physicians throughout the country adopted prepayment plans with

\textsuperscript{13}Rosen (1983), p.82.
the primary concern of ensuring payment. Most of the earlier plans offered full coverage without deductible for medical services up to a certain limit.\(^{14}\) Because these contracts were intended to keep benefits low in amount and short in duration, subscribers were exposed to unexpected major medical expenses. Despite this undesirable feature, the spread of health insurance was swift. Within the decade of 1941-1950, the percentage of population enrolled for hospital benefits increased from 12.4\% to 50.7\%.\(^ {15}\)

**After 1950**

During the latter half of the twentieth century, infectious diseases were no longer a major threat towards public health. As more people survived to old age, the prevalence of chronic diseases became more alarming. In 1900 diseases of the heart and cancers together accounted for only 15.1\% of all deaths.\(^ {16}\) Over the period 1900-1950, mortality rates of the two increased by 2.2 and 1.7 times, respectively [see Figure 1.8]. By the year 1950, 54.1\% of all deaths were due to these diseases and the share remained over 50\% since. During the 1950s and early 1960s the rising trend of life-expectancy was hindered by high death rates resulting from these two groups of diseases. The upward trend was not resumed until heart disease mortality began to decline sharply in the late 1960s [see Figures 1.4 and 1.8]. Between 1965 and 2001, heart disease mortality fell by 54.3\%. This coincided with the advent of many innovations used in cardiovascular treatment. Evidence suggests that these innovations had made a significant contribution to

---

\(^{14}\)Encouraged by the success of the prepayment associations, insurance companies also offered similar benefit patterns. The major difference is that insurance companies offered benefits in cash, whereas the prepayment associations offered benefits in services. Readers are referred to Faulkner (1960) for further details on the early health insurance plans.

\(^{15}\)Source: Reed (1965).

\(^{16}\)“Diseases of the heart” include ischemic (or coronary) heart diseases and other form of heart diseases.
the decline in mortality. Medical researchers Goldman and Cook (1984) reported that about 40% of the decline in coronary heart disease mortality over the period 1968-1976 can be attributed to life-saving innovations. More recently, Hunink, Goldman and Torsteson (1997) found that nearly half of the decline in coronary heart disease mortality during 1980-1990 was brought about by better medical treatment. Their results are supported by the work of Cutler, McClellan and Newhouse (1999), which found that 55% of the reduction in mortality from heart attacks during the period 1975-1995 can be attributed to improvement in medical treatment, especially the use of new pharmaceuticals.

As for cancers, despite the rising trend in cancer mortality, the probability of surviving the disease has been increasing during the post-war decades. The five-year relative survival rate for all forms of cancer increased from 35% to 62.7% over the period 1950-1995.\textsuperscript{17} Again, the march of science has made a significant contribution. Lichtenberg (2004) found that new drugs that emerged after 1970 could account for 50-60% of the increase in age-adjusted survival rates in the first six years after diagnosis. All this evidence suggests that technological progress in medical treatment played an important role in saving lives during the second half of the twentieth century.

Meanwhile, the increasing prevalence of chronic diseases and the rapid growth in medical costs raised concerns on the financial costs of prolonged illnesses. Since 1948 insurance companies began to offer a new type of plan, known as major medical expense policy, which provided coverage on the especially high costs of major and prolonged illnesses. The plan could be used either as a supplement for basic coverage, or as a separate comprehensive program. Conceivably, these policies

\textsuperscript{17}Lichtenberg (2004) argued that the increase in cancer mortality was largely the result of rising incidence.
would expose the insurers to tremendous liabilities. To ease these burdens, an initial deductible and a coinsurance provision were added so that the subscribers would bear some part of the costs. Over the period 1960-1986, the percentage of population protected by major medical expense policies increased from 17.6% to 65.8%.\footnote{Source: U.S. Bureau of Census, Statistical Abstract of the U.S., various issues.}

Another major development in the health insurance industry was the introduction of Medicare and Medicaid programs in 1966. The current study focuses on Medicare which targets primarily the elderly (over 65 years of age) population. One problem of the Medicare program is that it fails to cover very long hospital stays, which are highly expensive.\footnote{A brief description on the risk-sharing structure of Medicare can be found in the Appendix.} This exposes the beneficiaries to substantial medical expenditure risk. In response, over 60% of Medicare beneficiaries purchased supplementary private insurance in 2001.

\section{1.3 The Model}

\subsection{1.3.1 The Environment}

The model economy is composed of overlapping generations. In each period, a continuum of \textit{ex ante} identical agents is born. The size of cohort is growing at a constant rate $\gamma > 0$. Each agent begins his life with identical preferences, same level of health ($\overline{h}$) and zero wealth. In each period, each agent faces a positive probability of dying. The maximum age that one can live to is age $J$. Starting from age 0, an agent works until the exogenously given retirement age $I (< J)$ is reached. During the working years, an age-$j$ agent is endowed with $e_j$ units of effective labor which he supplies inelastically to the market. An individual supplies no labor when retired; i.e., $e_j = 0$ for $j = I + 1, \ldots, J$. There is also a
set of initial old agents. An agent who is of age \( j \geq 1 \) at time 0 is said to be of generation \(-j\). The health stock of these initial old agents at time 0 is again given by \( \bar{h} \).

In this economy there are two commodities: a consumption good and medical care. The former is produced by a neoclassical production function which will be described later. Each unit of consumption good can be transformed into \( \frac{1}{p} \) units of medical care. All medical care is used to produce new units of health via the production function \( i : \mathbb{R}_+ \to \mathbb{R}_+ \). The function is assumed to be twice continuously differentiable, strictly increasing and strictly concave. The accumulation process of health is given by

\[
h' = i(m) + (1 - \delta_h) h + \varepsilon, \tag{1.1}
\]

where \( h \) denotes health status at the beginning of the current age, \( h' \) is health status at the end of the current age, \( \delta_h \in (0, 1) \) is the depreciation rate of health and \( \varepsilon \) is a health shock. In each period, an agent faces an idiosyncratic health shock, \( \varepsilon \), drawn from a finite set \( \mathcal{E} = \{\varepsilon_1, \ldots, \varepsilon_S\} \). The severeness of the health shock is ranked according to \( \varepsilon_1 = 0 > \varepsilon_2 > \ldots > \varepsilon_S \). This shock is assumed to be independently distributed over time and across agents. In addition, the probability distribution of the shock is assumed to be age-dependent. Specifically, the probability of drawing \( \varepsilon \in \mathcal{E} \) at age \( j \) is denoted by \( \pi_j(\varepsilon) \), with \( \sum_{\varepsilon \in \mathcal{E}} \pi_j(\varepsilon) = 1 \) for all \( j \).

Conditional on being alive at the current age with end-of-period health status \( h' \), the probability of surviving to the next period is \( \Phi(h') \). The function \( \Phi : \mathbb{R} \to [0, 1] \) is made up of two parts: (i) for \( h' > 0 \), \( \Phi(h') \) is twice continuously differentiable, strictly increasing and strictly concave. This assumption is motivated by the fact that chronic diseases are much more likely to occur among the elderly. In the quantitative analysis, only chronic diseases are considered.
differentiable, strictly increasing, and satisfies \( \lim_{h' \to \infty} \Phi(h') = 1 \), (ii) for \( h' \leq 0 \), \( \Phi(h') = 0 \). The latter means death is certain when health falls below zero. Moreover, it is assumed that the composite function \( \tilde{\Phi}(m) \equiv \Phi[i(m) + (1 - \delta_k)h + \varepsilon] \) is strictly concave.

Utility is zero when deceased. The period utility function for a living agent is given by \( U(c) \), where \( c \) denotes current consumption. The utility function \( U : \mathbb{R}_+ \to \mathbb{R}_{++} \) satisfies all the usual assumptions and strict positivity; i.e., \( U(c) > 0 \) for \( c > 0 \). The last assumption is added to ensure that being alive with positive consumption is always preferable to being dead.

The consumption good is produced according to a Cobb-Douglas production function,

\[
y = Ak^\alpha l^{1-\alpha}, \quad \alpha \in (0, 1),
\]

where \( y \) denotes aggregate output, \( k \) denotes aggregate physical capital, \( A \) represents total factor productivity, and \( l \) is the aggregate labor input. The stock of physical capital accumulates according to

\[
k' = a + (1 - \delta_k)k,
\]

with the initial level of capital, \( k_0 \), given. The variable \( a \) represents gross investment, and \( \delta_k \in (0, 1) \) is the depreciation rate of physical capital.

1.3.2 The Benchmark Economy in 1950

Market Structure

In the framework described above, agents face two types of uncertainty: uncertainty concerning their end-of-period health and uncertainty with respect to the
length of their lifetime. This section describes an economy in which agents can
insure against the health risk but not the mortality risk. In this economy, there is
a private insurance market in which reimbursement health insurance contract is
being traded. Private annuities are missing, so agents cannot protect themselves
against the uncertainty in consumption brought by an uncertain length of lifetime.
In terms of investment opportunities, physical capital is the only channel for in-
vestment and agents are not allowed to borrow. In the quantitative analysis, this
model is used to represent the U.S. economy in 1950, a time before the Medicare
program is implemented. Hence, public health insurance is not considered in here.

In the private insurance market, insurance companies cannot observe the
health status of their customers. All these companies can observe are their cus-
tomers’ medical expenses. As a result, insurance benefits are paid as reimburse-
ment on the actual expenses, which are controlled by the insured. It is assumed
that a standard, perfectly divisible contract is being traded. This means agents
can buy any positive amount, \( \bar{x} \geq 0 \), of this contract. The health insurance con-
tract is characterized by a premium rate \( \pi_p \) and a piecewise linear reimbursement
function \( \Theta : \mathbb{R}_+ \rightarrow \mathbb{R}_+ \). The latter is a function of the actual expenses incurred
by the insured, \( pm \), and takes the following form

\[
\Theta(pm) = \begin{cases} 
0 & \text{if } pm \leq d \\
\kappa (pm - d) & \text{if } pm \in [d, l] \\
L & \text{if } pm > l, 
\end{cases}
\]

where \( d \geq 0 \) is the deductible, \( \kappa \in (0, 1) \) is the coinsurance rate and \( L \) is the
maximum amount that the insurance company is willing to pay. The reimburse-
ment function is assumed to be continuous on \( \mathbb{R}_+ \), so \( L = \kappa (l - d) \). Suppose
an agent with a medical bill $pm$ purchased $\bar{x} > 0$ units of insurance. The agent would split the medical bill into $\bar{x}$ shares and subject each portion to a separate reimbursement schedule. His out-of-pocket expenditures are then given by

\[
pm - \bar{x} \Theta(pm/\bar{x}) = \begin{cases} 
\pm & \text{if } pm/\bar{x} \leq d \\
(1 - \kappa) \pm + \bar{x} \kappa d & \text{if } pm/\bar{x} \in [d,L] \\
\pm - \bar{x} L & \text{if } pm/\bar{x} \geq L.
\end{cases}
\]

In the absence of an annuity market, agents who die prematurely (before age $J$) will leave behind their wealth as accidental bequests. These are distributed to all of the survivors in a lump-sum fashion.

The sequence of events is as follows: At the beginning of each period, an agent chooses the amount of private health insurance and a set of state-contingent decision rules for consumption, medical care and savings. All these decisions are made before the current health shock is realized. After it is realized, commodities specified by the decision rules are delivered and reimbursement from the insurance is made. Before the end of the period, the agent is informed whether or not he will survive in the next period.

**Individual’s Problem**

Consider an age-$j$ agent who enters the period with characteristics $\theta = (h, s_{-1})$, where $s_{-1}$ is previous savings in terms of physical capital. Let $w$ be the market wage rate for effective unit of labor, $r$ be the (gross) return from physical capital and $b$ be the per-capita amount of accidental bequests. The agent chooses the amount of insurance and a set of state-contingent decision rules, denoted by
\( \varphi = \{ \bar{x}_j(\theta), c_j(\theta, \varepsilon), m_j(\theta, \varepsilon), s_j(\theta, \varepsilon) | \varepsilon \in \mathcal{E} \} \), to solve the following problem

\[
E^j \left[ V^j(\theta, \varepsilon) \right] = \max_{\varphi} E^j \left\{ U[c_j(\theta, \varepsilon)] + \beta \Phi[h_{j+1}(\theta, \varepsilon)] E^{j+1} \left[ V^{j+1}(\theta', \varepsilon') \right] \right\}
\]

subject to

\[
c_j(\theta, \varepsilon) + pm_j(\theta, \varepsilon) + s_j(\theta, \varepsilon) = we_j + rs_{-1} + b + \bar{x}_j(\theta) \left\{ \Theta \left[ \frac{pm_j(\theta, \varepsilon)}{\bar{x}_j(\theta)} \right] - \pi_p \right\},
\]

\[
h_{j+1}(\theta, \varepsilon) = i[m_j(\theta, \varepsilon)] + (1 - \delta_h) h + \varepsilon,
\]

\[
s_j(\theta, \varepsilon) \geq 0,
\]

for \( \varepsilon \in \mathcal{E} \) and \( \bar{x}_j(\theta) \geq 0 \).

Denote by \( z^j = (\varepsilon^0, ..., \varepsilon^j) \) a history of health shock up to age \( j \) and let \( Z^j \) be the set of all possible \( z^j \). The notation \( z^{-1} \) is used to denote the empty past history of a newborn agent. The following lemma establishes that the agent’s decisions for consumption, medical spending, saving and insurance can be expressed solely as a function of the history of health shocks.

**Lemma** The solution of (P1.1) can be expressed as \( c_j(z^j), m_j(z^j), s_j(z^j) \) and \( \bar{x}_j(z^{j-1}) \).

**Proof.** See Appendix.  

**Competitive Equilibrium**

This section describes a competitive equilibrium for a stationary environment in which the life-cycle patterns of individual consumption, medical spending and savings are time-invariant. It follows that all the aggregate variables and the population structure are time-invariant in this environment.
Let \( N(z^j) \) denote the measure of all age-\( j \) agents with shock history \( z^j \). The population measure evolves according to

\[
N(z^j) = \pi_j(\varepsilon) \Phi[h_j(z^{j-1})] N(z^{j-1}), \quad \text{for } j = 1, \ldots, J, \tag{1.4a}
\]

with \( z^j = (z^{j-1}, \varepsilon) \), \( \Phi[h_j(z^{j-1})] \equiv \Phi[h_j(z^{j-1})] / (1 + \gamma) \), and

\[
N(z^0) = \pi_0(\varepsilon), \tag{1.4b}
\]

for \( z^0 = \varepsilon \in \mathcal{E} \). The size of the total population is given by

\[
N = \sum_{j=0}^{J} \sum_{z^j} N(z^j).
\]

Among those age-\( j \) agents with shock history \( z^j \), a fraction \( \{1 - \Phi[h_{j+1}(z^j)]\} \) will be deceased in the next period. Each deceased agent will leave behind an amount \( rs_j(z^j) \). The amount of bequest received by each survivor is thus given by

\[
b = \frac{1}{N} \sum_{j=0}^{J-1} \sum_{z^j} N(z^j) \left\{1 - \Phi[h_{j+1}(z^j)]\right\} rs_j(z^j) = 0. \tag{1.5}
\]

All the savings are channelled to the capital market. The capital market clears when aggregate demand equals aggregate supply; i.e.,

\[
k = \sum_{j=0}^{J} \sum_{z^j} N(z^j) s_j(z^j). \tag{1.6}
\]

In the consumption good sector, firms hire labor and rent physical capital from the factor markets to produce output according to (1.2). Both goods and factor
markets are assumed to be perfectly competitive, hence

\[ r = \alpha A k^{\alpha - 1} l^{1-\alpha} + 1 - \delta_k, \quad (1.7) \]

\[ w = (1 - \alpha) A k^{\alpha} l^{-\alpha}. \quad (1.8) \]

The consumption good market clears when

\[
\sum_{j=0}^{J} \sum_{z^j} N(z^j) \left[ c_j(z^j) + p m_j(z^j) \right] + \delta_k k = A k^{\alpha} l^{1-\alpha}. \quad (1.9)
\]

The labor market clears when the following holds,

\[ l = \sum_{j=0}^{I} \sum_{z^j} N(z^j) e_j. \quad (1.10) \]

The private insurance market is assumed to be perfectly competitive. The insurance premium, \( \pi_p \), is endogenously determined by the market-clearing condition:

\[
\sum_{j=0}^{J} \sum_{z^{j-1}} \sum_{\varepsilon \in \mathcal{E}} N(z^{j-1}, \varepsilon) \tilde{x}_j(z^{j-1}) \left\{ \Theta \left[ \frac{p m_j(z^{j-1}, \varepsilon)}{\tilde{x}_j(z^{j-1})} \right] - \pi_p \right\} = 0. \quad (1.11)
\]

A competitive equilibrium for this economy is defined below:

**Definition** A competitive equilibrium for this economy consists of allocations

\[ \varphi = \{ c_j(z^{j-1}, \varepsilon), m_j(z^{j-1}, \varepsilon), s_j(z^{j-1}, \varepsilon), \tilde{x}_j(z^{j-1}) \mid z^{j-1} \in \mathbb{Z}^{j-1}, \varepsilon \in \mathcal{E} \}_{j=0}^{J} \], a set of prices \( \{ r, w, \pi_p \} \), factor inputs \( \{ k, l \} \), population measures \( \{ N(z^j) \mid z^j \in \mathbb{Z}^j \}_{j=0}^{J} \), and per-capita bequest \( b \) such that

1. Given the prices and bequest, \( \varphi \) solves the agent’s problem.

2. Factor inputs and prices, \( \{ k, l, r, w \} \), satisfy (1.7) and (1.8).
3. The population measures evolve according to (1.4a) and (1.4b).

4. Per-capita bequest, $b$, is given by (1.5).

5. All markets clear; i.e., (1.6), (1.9), (1.10) and (1.11) hold.

1.3.3 The Benchmark Economy in 2001

Market Structure

Moving forward across time, the year is now 2001. The only difference between 1950 and 2001 in terms of insurance opportunity is that public health insurance programs, such as Medicare and Medicaid, are present in 2001. To capture this, a public health insurance program similar to Medicare is introduced into the economy described above. In the quantitative exercise, this version of the benchmark economy is calibrated to mimic the U.S. economy in 2001.

Under the public health insurance program, all retirees are automatically enrolled at zero cost. The program offers a reimbursement schedule $\Theta_g : \mathbb{R}_+ \to \mathbb{R}_+$ that is similar to private insurance,

$$
\Theta_g(pm) = \begin{cases} 0 & \text{if } pm \leq d_g \\ \kappa_g(pm - d_g) & \text{if } pm \in [d_g, l_g] \\ L_g & \text{if } pm > l_g. \end{cases}
$$

Notice that for $pm > 0$, the program covers only part of the medical expenses. This creates an incentive for the retirees to purchase supplementary insurance from the private health insurance market to cover the differences, $pm - \Theta_g(pm)$, that are not insured by the public health insurance program. The public insurance program is financed by a payroll tax $\tau \in (0, 1)$ imposed on the working population.
For the working agents, private health insurance is the only source of insurance. Expenditures on private insurance can be deducted from the payroll tax. The rest of the economy is the same as that described in section 1.3.2.

**Individual’s Problem**

First, consider a retired agent of age $j > I$ who begins the period with characteristics $\theta = (h, s_{-1})$. His task is to choose $\varphi = \{\tilde{x}_j(\theta), c_j(\theta, \varepsilon), m_j(\theta, \varepsilon), s_j(\theta, \varepsilon) | \varepsilon \in \mathcal{E}\}$ in order to maximize the expected value of his remaining life,

$$
E^j \left[ V^j(\theta, \varepsilon) \right] = \max_{\varphi} \left\{ U \left[ c_j(\theta, \varepsilon) \right] + \beta \Phi \left[ h_{j+1}(\theta, \varepsilon) \right] E^{j+1} \left[ V^{j+1}(\theta', \varepsilon') \right] \right\}
$$

subject to

$$
c_j(\theta, \varepsilon) + v_j(\theta, \varepsilon) + s_j(\theta, \varepsilon) = r s_{-1} + b + \tilde{x}_j(\theta) \left\{ \Theta \left[ \frac{v_j(\theta, \varepsilon)}{\tilde{x}_j(\theta)} \right] - \pi_p \right\},
$$

$$
v_j(\theta, \varepsilon) = p m_j(\theta, \varepsilon) - \Theta_g \left[ p m_j(\theta, \varepsilon) \right],
$$

$$
h_{j+1}(\theta, \varepsilon) = i \left[ m_j(\theta, \varepsilon) \right] + (1 - \delta_h) h + \varepsilon,
$$

$$s_j(\theta, \varepsilon) \geq 0,$$

for $\varepsilon \in \mathcal{E}$ and $\tilde{x}_j(\theta) \geq 0$. The variable $v_j(\theta, \varepsilon)$ represents the amount of medical expenditures that are not covered by public insurance but potentially covered by supplementary private insurance.

Next, consider an agent of working age $(j \leq I)$ with characteristics $\theta = (h, s_{-1})$. The problem faced by this agent is to choose a set of decision rules on consumption, medical services, savings and insurance to maximize his expected lifetime utility,
\[
E^j \left[ V^j (\theta, \varepsilon) \right] = \max_{\varphi} E^j \left\{ U (c) + \beta \Phi [h_{j+1} (\theta, \varepsilon)] E^{j+1} \left[ V^{j+1} (\theta', \varepsilon') \right] \right\} \quad (P1.2')
\]

subject to

\[
c_j (\theta, \varepsilon) + pm_j (\theta, \varepsilon) + s_j (\theta, \varepsilon) = (1 - \tau) \left[ we_j - \pi_p \bar{x}_j (\theta) \right] + rs_{-1} + b + \bar{x}_j (\theta) \Theta \left[ \frac{pm_j (\theta, \varepsilon)}{\bar{x}_j (\theta)} \right],
\]

\[
h_{j+1} (\theta, \varepsilon) = i \left[ m_j (\theta, \varepsilon) \right] + (1 - \delta_h) h + \varepsilon,
\]

\[
s_j (\theta, \varepsilon) \geq 0,
\]

for \( \varepsilon \in E \) and \( \bar{x}_j (\theta) \geq 0 \). As in the benchmark economy of 1950, all the decision rules can be expressed in terms of the shock history.

**Competitive Equilibrium**

In this economy, the only function of the government is to maintain the public insurance program and balance its budget in every period. The government’s budget constraint is given by

\[
\tau \sum_{j=0}^{\infty} \sum_{z^{j-1}} \sum_{\varepsilon \in E} N \left( z^{j-1}, \varepsilon \right) \left[ we_j - \pi_p \bar{x}_j (z^{j-1}) \right] = \sum_{j=I+1}^{\infty} \sum_{z^j} N \left( z^j \right) \Theta_g \left[ pm_j (z^j) \right].
\]

On the left-hand side is the total payroll tax collected less the tax-deductible expenditures on private health insurance. The right-hand side gives the total reimbursement received by the beneficiaries of the public insurance program.\(^{21}\)

The rest of the economy is the same as that described in previous section.

A competitive equilibrium for this economy is defined as follow:

\(^{21}\)In the quantitative analysis, the tax rate \( \tau \) is taken as exogenous given, while the reimbursement function, \( \Theta_g (\cdot) \), is endogenously determined.
Definition  A competitive equilibrium for this economy consists of allocations
\[ \varphi = \{ c_j(z^{j-1}, \varepsilon), m_j(z^{j-1}, \varepsilon), s_j(z^{j-1}, \varepsilon), x_j(z^{j-1}), z^{j-1} \in Z^{j-1}, \varepsilon \in E \}_{j=0}^J, \]  

a set of prices \( \{r, w, \pi_p\} \), factor inputs \( \{k, l\} \), population measures \( \{N(z^j) | z^j \in Z^j \}_{j=0}^J \), income tax rate \( \tau \) and per-capita bequest \( b \) such that

1. Given the prices and bequest, \( \varphi \) solves agents’ problem.

2. Factor inputs and prices, \( \{k, l, r, w\} \), satisfy (1.7) and (1.8).

3. The population measures evolve according to (1.4a) and (1.4b).

4. Per-capita bequest, \( b \), is given by (1.5).

5. The government’s budget is balanced in each period, i.e. (1.12) holds.

6. All markets clear; i.e., (1.6), (1.9), (1.10) and (1.11) hold.

1.3.4 The Economy with No Insurance

This section describes an economy in which agents are cut off from all insurance markets. To insure themselves against consumption fluctuations brought by the health shock and mortality risk, the agents can only adjust their holdings of physical capital. Again, agents are not allowed to borrow. The agent’s problem is to choose a set of state-contingent decision rules, \( \varphi = \{ c_j(\theta, \varepsilon), m_j(\theta, \varepsilon), s_j(\theta, \varepsilon) | \varepsilon \in E \} \), in order to maximize the expected value for the remaining lifetime. Formally, this is written as

\[
E^j [V^j(\theta, \varepsilon)] = \max_{\varphi} E^j \left\{ U [c_j(\theta, \varepsilon)] + \beta \Phi [h_{j+1}(\theta, \varepsilon)] E^{j+1} [V^{j+1}(\theta', \varepsilon')] \right\}
\]

subject to

\[ c_j(\theta, \varepsilon) + pm_j(\theta, \varepsilon) + s_j(\theta, \varepsilon) = w e_j + rs_{-1} + b, \]
\[ h_{j+1}(\theta, \varepsilon) = i[m_j(\theta, \varepsilon)] + (1 - \delta_h) h + \varepsilon, \]

for \( \varepsilon \in \mathcal{E} \). Since this is just a special case of the benchmark model, a competitive equilibrium for this economy can be defined in a similar fashion.

### 1.4 Quantitative Analysis

The objective of the calibration exercise is to construct two steady states, based on the benchmark economies, to represent the U.S. economy in 1950 and 2001. Most of the parameters in the model can be readily obtained from the data. For those that cannot, a minimization procedure is used to determine their values. Table 1.3 summarizes all the parameters used in the baseline calibration.

**Demography** In the model economy, one period takes 5 years. Individuals are assumed to be economically active at age 25. Thus “age 0” in the model refers to the 25-29 age group. Take \( I = 7 \) and \( J = 15 \) so that retirement begins at age 65 and the oldest age group is 100-104. Over the period 1950-2001, the U.S. population grew at an average rate of 1.25% per year. In the model, the parameter \( \gamma \) corresponds to the 5-year population growth rate. Hence, \( \gamma = (1.0125)^5 - 1 = 0.064 \).

The survival probability function, \( \Phi(\cdot) \), is assumed to take the form of the cumulative distribution function of a Weibull distribution:

\[ \Phi(h) = 1 - \exp(-\psi h^\theta), \quad \text{for } h \geq 0, \]

with \( \psi > 0 \) and \( \theta > 0 \). Both \( \psi \) and \( \theta \) are determined by the minimization procedure. The initial level of health at age 0, \( \bar{h} \), is assumed to be constant over time and is normalized to 100.
Preferences and Earnings The period utility function takes the standard CRRA form; i.e., \( U(c) = \frac{c^{1-\sigma}}{1-\sigma} \). Conventionally, the coefficient of relative risk aversion is no less than one. But then utility might be negative for some positive values of consumption. Thus, the coefficient must be strictly less than one in order to ensure strict positivity of utility. In the baseline calibration, the coefficient of relative risk aversion is set to 0.97. Section 1.5.3 reports how the main findings are affected when this coefficient changes. The annual subjective discount factor is taken to be 0.98, so \( \beta = (0.98)^5 = 0.904 \). The age-specific effective units of labor, \( \{e_j\}_{j=0}^7 \), are calibrated using data on money earnings in 1950 and 2001.\(^{22}\) The parameter values are reported in Table 1.3b.

Production Technology In the production function for the consumption good, labor’s share of income is fixed at 0.67, so \( \alpha = 0.33 \). The annual depreciation rate for physical capital is taken to be 10\%. Hence, \( \delta_k = 1 - (1 - 0.1)^5 = 0.410 \).

Morbidity At each age, there are two possible values of health shock, \( \{\varepsilon_1, \varepsilon_2\} \). The first one corresponds to the state of being “healthy”, or one without any negative health shock so that \( \varepsilon_1 = 0 \). The second one corresponds to the state of being “sick”. In the current context, being “sick” means suffering from either coronary heart disease (CHD) or cancers of any form.\(^{23}\) The probabilities of the two states are computed using the incidence rates of these diseases. Figures 1.9 and 1.10 plot the average annual age-specific incidence rates for CHD and cancers.

\(^{22}\)For 1950, data on wage and salary incomes obtained from the IPUMS general sample are used. For 2001, mean earnings by age reported in the Current Population Surveys are used. The labor endowment of those aged 25-29 in 1950 is normalized to unity. The effective units of labor for all other demographic groups are then derived by the ratio of mean earnings of that group to the mean earnings of the reference group.

\(^{23}\)The importance of these diseases are discussed in section 1.2.2.
respectively.\textsuperscript{24} Let $q_j^h$ be the annual incidence rate for CHD at age $j$ and $q_j^c$ be the corresponding rate for cancers. In the model, being “healthy” means not suffering from CHD \textit{and} not suffering from cancers for a period of five years. Hence, the probability of being “healthy” at age $j$ is

$$\pi_j(\varepsilon_1) = [\left(1 - q_j^h\right) \left(1 - q_j^c\right)]^5.$$ 

The probability of being “sick” is $\pi_j(\varepsilon_2) = 1 - \pi_j(\varepsilon_1)$. The values of these probabilities are listed in Table 1.3c. The magnitude of the negative health shock ($\varepsilon_2$) is determined by the minimization procedure.

\textbf{Medical Technology and Prices} \ The production function for health at time $t$ is specified as follows:

$$i_t(m) = \eta_t m^\xi,$$  \hspace{1cm} (1.13)

where $\eta_t > 0$ and $\xi \in (0, 1)$. The productivity of medical care at time $t$ is captured by $\eta_t$, while the cost of medical care is $p_t$. Both are exogenously given in the current model. Intuitively, medical care includes such commodities as a hospital stay or a physician visit, which are inputs into the production for health. Innovations in medical treatment are likely to raise the price of these inputs because they are now more productive, and lower the price of output (i.e., health). Given the specific functional form in (1.13), the price of health in terms of consumption is given by

$$\frac{p_t}{i_t'(m)} = \frac{p_t}{\eta_t \xi} m^{1-\xi}.$$ 

\textsuperscript{24}Despite their alarming prevalence during the post-war era, nationwide data concerning the incidence of CHD and cancers are not available until the 1980s. Source: (i) CHD, The Atherosclerosis Risk in Communities Study (ARIC), <http://www.cscce.unc.edu/aric>. (ii) Cancers, National Cancer Institute, \textit{SEER Cancer Statistics Review}, various issues.
Holding other things constant, the price of health would fall if $\eta_t$ grows faster than $p_t$.

In the health economics literature, attempts to quantify the pace of technological progress in medical treatment are sparse. In a recent study, Lichtenberg and Virabhak (2002) estimated the rate of technological progress embodied in different vintages of drugs. The estimated growth rate ranged from 3.1% to 4.2% depending on the measure of health.$^{25}$ In the calibration, the average annual growth rate of $\eta_t$ is taken to be 3.4%. Hence, $\eta_{2001} = (1.034)^{51} \eta_{1950}$. As for medical inflation, it is measured by the rate of change of the medical care component in the CPI relative to the general price level. Over the period 1950-2001, this component inflated at an average annual rate of 2.1% relative to the GDP deflator.$^{26}$ The cost of medical care in 2001 is determined by $p_{2001} = (1.021)^{51} p_{1950}$. In the baseline calibration, the price of medical care in 1950 is normalized to 10. In section 5.3, it is shown that the main findings are not sensitive to the choice of $p_{1950}$. The growth rates of $\eta$ and $p$, together with the functional form in (1.13), imply that the price of health declined at an average annual rate of 1.26% over the period 1950-2001.$^{27}$

$^{25}$There are two additional reasons why these results are suitable for the current study. First, Lichtenberg and Virabhak considered drugs that are approved over an extensive time period, from 1939 to 1998. Second, these estimates are based on a “dynamic” equation in which the pre-treatment (or beginning-of-period) health is controlled for. This “dynamic” equation is analogous to equation (1.1) in the current work.

$^{26}$Admittedly the medical CPI may not be the best measure of medical inflation. A detailed account on the construction and shortcomings of this subindex can be found in Berndt et al (2001). This is used because there is no alternative measure for general medical inflation that covers the period in question.

$^{27}$Cutler, McClellan, Newhouse and Remler (1998) estimated the quality-adjusted cost of treating heart attack over the period 1983-1994. Similar to the current study, they use length of life as the measure of health. According to their benchmark estimate, the quality-adjusted cost fell at an average annual rate of 1.1%.
Health Insurance  In the calibration, all hypothetical insurance contracts are assumed to be major medical expense contracts. During the 1950s, a typical contract of this sort had an initial deductible ranging from $50 to $500 (in current dollars) and a coinsurance clause that required the insured to pay 20% to 25% of the expenses above the deductible. In 2001, a typical contract in the private insurance market had the same range of coinsurance rate and a deductible that varied between $250 and $2500 (in current dollars). In the calibration, all private insurance contracts have the same coinsurance rate of 75%; i.e., $\kappa_{1950} = \kappa_{2001} = 0.75$. In the 1950 and 2001 steady states, the deductible are taken to be $100 and $250 (in current dollars), respectively.

As for Medicare in 2001, the cost-sharing structure is simplified so that it shares the same structure as private insurance. Specifically, it involves a deductible of $892 and a coinsurance rate of 75%; i.e., $\kappa_{g} = 0.75$.

The Remaining Parameters  Up to this point, the values of nine parameters are not yet determined. These include parameters in the survival probability function ($\psi, \theta$) and the health production function ($\eta_{1950}, \xi$), the depreciation rate of health capital ($\delta_h$), the magnitude of the negative health shock ($\varepsilon_2$), and the maximum reimbursement levels for health insurance ($L_{1950}, L_{2001}, L_g$). A minimization procedure is used to determine these parameter values. The idea is to choose these values so that the model could match, as closely as possible, the nine real-world statistics listed in Table 1.4. Formally, let $\chi$ be a (column) vector of parameters, and $S$ be a (column) vector of selected real-world statistics. Given $\chi$, the model could yield a prediction on $S$, denoted by $\hat{S}(\chi)$. The minimization

\[28\text{Source: Reed (1965).}\]
procedure then involves solving the following problem:

$$\min_{\chi} \left[ \tilde{S}(\chi) - S \right]^T \left[ \tilde{S}(\chi) - S \right].$$

### 1.5 Findings

This section is organized into three parts. The first part documents the main findings obtained from the benchmark economies. These findings are summarized in Table 1.5. The second part reports the results obtained when the parameterization described in section 4 is imposed on the economy with no insurance. The same subsection also answers the question that motivates the current work, namely “How much increase in life expectancy and medical spending can be attributed to technological progress in medical treatment and rising income?” The last part of this section illustrates how the baseline results are affected when some of the parameters vary.

#### 1.5.1 Benchmark Economy

**Medical Expenditures** Under the baseline parameterization, the model is able to capture the expansion in medical spending between 1950 and 2001. In the model economy, the share of medical spending in GDP increases from 3.7% to 12.4%. In terms of personal consumption expenditures (PCE), the share of medical spending increases from 4.8% to 16.6%. Between the two steady states, real per-capita medical spending increase by a factor of 9.13. The actual growth factor between 1950 and 2000 is 8.81. In terms of average annual growth rate, the predicted value is 4.43%, while the observed rate is 4.44%. As a reference, real per-capita GDP increases at an average annual rate of 2.0% in the model economy and the actual
rate is 2.1%. In the 1950 steady state, 68.2% of the total medical spending is paid directly by the consumers. In the 2001 steady state, out-of-pocket expenditures accounted for 24.7% of the total spending, while Medicare accounted for 15.5%. This shows that there is a large expansion in third-party payments in the model economy.

**Life Expectancy** The model’s predictions on life expectancy are depicted in Figures 1.11 and 1.12. The underlying survival probabilities are shown in Figures 1.13 and 1.14. Under the baseline parameterization, life expectancy increases by 4.1 years at age 25, 4.0 years at age 45 and 3.1 years at age 65. When compared to the data, the model is able to explain 64.6% of the increase in life expectancy at age 25, 67.8% at age 45 and 79.5% at age 65. Various measures are devised to capture the overall changes in life expectancy. The results are reported in Table 1.6. In summary, the model can explain more than 60% of the increase in average life expectancy. The precise number depends on which measure is used.

**Life-cycle Medical Spending** This subsection examines the model’s predictions on the life-cycle patterns of medical spending. Figures 1.15 and 1.16 show the predicted and observed ratio of medical spending to income across various age groups. The first measure of average life expectancy is constructed as below:

\[
\Gamma_{1t} = \sum_{j=0}^{J} \Psi_{j,t} \Gamma_{j,t}, \quad \text{for } t = 1950, 2001,
\]

where \(\Psi_{j,t}\) is the population share of age-\(j\) agents at time \(t\) and \(\Gamma_{j,t}\) denotes life expectancy at age \(j\). The second measure is constructed by holding the population structure constant as in 1950. In practice, this involves replacing \(\Psi_{j,t}\) with \(\Psi_{j,1950}\) in the above expression. The third measure is constructed by holding the population structure constant as in 2001. The numbers reported in Table 6 are the differences between the two time periods. The actual numbers can be found in Table 1.7 panel A. These are the ratios of average medical spending for a particular age group to the average income of that group. For 1950,
64 and increases significantly afterwards. The current model is able to replicate these patterns. In the 1950 steady state, the ratio of medical spending to income before and after age 65 are 0.04 and 0.105, respectively. In the data, the corresponding figures are 0.037 and 0.117. In 2001, the predicted values are 0.123 and 0.635 while the observed values are 0.096 and 0.612. In the model, the mild increase before age 65 is largely the result of a low depreciation rate in health. The intuition is as follow. Since the probability of being sick when young is very low, the main reason why a young agent would invest in his health is to offset the depreciation. A low depreciation rate in health thus implies a low growth rate in medical spending when young. The model is able to yield reasonable predictions on the growth rate of medical spending by age. The predicted growth rate of medical spending among the elderly and the non-elderly are 5.48% and 4.26% per year, respectively. The actual annual growth rates over the period 1950-2000 are 5.41% and 3.81% [see Table 1.7 panel B].

One problem with the current model is that it predicts a sharp decline in spending among those over 75 years of age. The predicted life-cycle profiles of medical spending thus exhibit a hump-shaped pattern which is not observed in the data. This reflects a low demand for health among the oldest agents in the model. In general, these agents are poor in health and have a short expected lifespan. The former factor encourages investment in health while the latter suppresses it. The quantitative results suggest that the suppressing force dominates. In the current framework, health is demanded only because it extends life. Admittedly this ignores other attributes of health. Daily experience suggests that a person with poor health may not face any immediate risk of losing life but may suffer median incomes reported in *Census of Population: 1950*, vol. II, Table 139 are used. For 2000, data on median incomes are obtained from U.S. Census Bureau, Current Population Survey, <http://www.census.gov/hhes/www/income/income01/inctab7.html>.
from some form of disability. Recent studies on medical spending have shown that disability status is closely related to spending among the elderly. Focusing on the Medicare beneficiaries in 1989-1990, Cutler and Meara (2001) reported that an average person aged 85 or above spent almost $2,000 more on medical care than one aged 65-69. Most importantly, a large part of these differences can be explained by differences in disability status. Ignoring other attributes of health may be the reason why the model tends to underestimate the average spending for the elderly.

Life-cycle Consumption It is well documented that life-cycle consumption profiles are hump-shaped in nature. In a recent study, Fernandez-Villaverde and Krueger (2002) estimated the life-cycle consumption profiles using data from the Consumer Expenditures Survey over the period 1980-1998. They found that after adjusting for life-cycle changes in family size, non-durable consumption peaked at age 52 and was about 29% higher than that at age 25. The current model is able to yield similar hump-shaped patterns but the peaks are lower than that reported by Fernandez-Villaverde and Krueger. In the 1950 steady state, consumption reaches its peak at ages 45-49. The peak level is about 19% higher than that in ages 25-29. In the 2001 steady state, consumption peaks at ages 50-54. The ratio of peak consumption to that in ages 25-29 is around 1.18.

1.5.2 Economy with No Insurance

Table 1.8 compares the results obtained from the two economies under the same parameterization. Two observations can be made from these findings. First, in both 1950 and 2001 the economy with no insurance could yield similar (but

---

31 Consumption expenditures are the sum of spending on consumption good, out-of-pocket expenditures on medical care and spending on private health insurance.
slightly lower) levels of life expectancy with fewer medical spending. When all insurance opportunities are removed, the share of medical spending in GDP would be lowered by 1.6 percentage points in 1950 and 1.4 percentage points in 2001. In terms of real per-capita spending, these represent a 43% reduction in 1950 and a 11% reduction in 2001. These findings are consistent with the idea that people tend to spend more on medical services in the presence of reimbursement insurance. Second, and most importantly, a large increase in medical spending between 1950 and 2001 is observed even when all insurance opportunities are removed. In this case, the share of medical spending increases by 8.9 percentage points, comparing to the actual increase of 8.5 percentage points. Thus, the model suggests that the increase in medical spending during the latter half of the twentieth century is not driven by factors associated with insurance opportunities.\textsuperscript{32} Instead, the increase is driven by factors that are common in both economies, namely technological progress in medical treatment and rising incomes. With these two factors alone, the model is able to explain 63\% of the increase in life expectancy at age 25.

When both the price and productivity of medical care are kept constant; i.e., \( p_{1950} = p_{2001} \) and \( \eta_{1950} = \eta_{2001} \), the share of medical spending is merely 2.3\% in 2001. The resulting levels of life expectancy are almost identical to those in 1950.\textsuperscript{33} Thus the current model suggests that income growth alone is not enough to explain the observed changes in medical spending and life expectancy.

\textsuperscript{32}In Hall and Jones (2004), this proposition is taken as given. In the current work, this is formally derived from a dynamic general equilibrium model.

\textsuperscript{33}Similar results can be obtained when the experiment is performed in an economy with private insurance only. In this case, the share of medical spending in 2001 is 3.9\%, comparing to the benchmark result of 12.4\%. 
1.5.3 Robustness

In the first robustness check, different values of the coefficient of relative risk aversion \((\sigma)\) are considered.\(^34\) In general, a higher degree of risk aversion is associated with a larger demand for insurance and a higher level of medical spending (see Table 1.9a). The demand for insurance is measured by the per-capita amount of insurance purchased. Algebraically, this is given by

\[
\frac{1}{N} \sum_{j=0}^{J} \sum_{z^{-1}} \sum_{\varepsilon \in \mathcal{E}} N(z_j^{-1}, \varepsilon) \bar{x}_j(z_j^{-1}) .
\]

The findings are consistent with the idea that more risk averse agents would like to purchase more insurance. This in turns leads to a higher level of medical spending. An alternative way to assess the change in medical spending in each case is to re-calibrate the productivity of medical care in 1950 \((\eta_{1950})\) so as to match the observed share of medical spending in 1950. The results are shown in Table 1.9b. This exercise illustrates that by adjusting one or more of the undetermined parameters, similar increase in medical spending can be obtained for each value of \(\sigma\).

In the second robustness check, the price of medical care in 1950 \((p_{1950})\) is varied. Table 1.10 shows that the results on medical spending are not sensitive to the choice of \(p_{1950}\). For instance, the share of medical spending reduces from 3.7% to 3.4% when \(p_{1950}\) increases from 10 to 20.

\(^{34}\)In the 2001 economy, the maximum reimbursement of Medicare \((L_g)\) varies in each case in order to keep the tax rate constant. All other parameters are held constant.
1.6 Concluding Remarks

Two trends are observed during the latter half of the twentieth century. First, there is a persistent rising trend in medical spending. Second, there is a significant increase in life expectancy. The hypothesis examined in this paper is that the combination of technological improvements in medical treatment and rising incomes is the driving force behind these two trends. The backbone of the current analysis is a stochastic, multi-period overlapping-generations model with endogenous survival probability. Agents are ex post heterogeneous in terms of the realizations of health shocks. Two market structures with different insurance opportunities are imposed on this environment. In the first economy, public and/or private health insurance are available via which the consumers can insure against the health shock. In the second economy, agents can only self-insure.

Two things are learned from the quantitative analysis. First, in the presence of technological progress in medical treatment and rising incomes, large increases in medical spending can be obtained in both the benchmark economy and the economy with no insurance. Hence, unlike the previous literature, the current model suggests that the rapid growth in medical spending is not due to factors associated with market structures or insurance opportunities. Second, in the current framework technological progress in medical treatment and rising incomes can explain all the increase in medical spending and more than 60% of the increase in average life expectancy during the second half of the twentieth century. This suggests that the rapid growth in medical spending reflects optimal responses to changes underlying the production and accumulation of health.

In mainstream macroeconomic studies, an agent’s planning horizon is both fixed and predetermined. The current work is an initial attempt to explore the
macroeconomic implications of an endogenous and stochastic planning horizon. Growth in medical spending is just one of the many phenomena associated with a longer lifespan. Intuitively, increases in longevity would have an impact on other life-cycle decisions such as savings, educational choices and retirement. The framework presented in this chapter can be extended easily to study these issues.
1.7 Appendix

1.7.1 Medical Spending by Age, 1950-2000

Meara, White and Cutler (2004) estimated the per-person medical spending across age groups for the years 1963, 1970, 1977, 1987, 1996 and 2000. The results, tabulated in their technical appendix, are duplicated in Table A1. This section describes how the estimates for the year 1950 are obtained in this paper. First, it is assumed that over the period 1950-1963, per-person spending in any specific age group had been growing at the same age-specific rate as in 1963-1977. Formally, let \( m_{j,t} \) denote the average spending at age \( j \) in year \( t \). The age-specific, average annual growth rates \( \gamma_j \) between 1963 and 1977 are defined by

\[
(1 + \gamma_j)^{14} = \frac{m_{j,1977} - m_{j,1963}}{m_{j,1963}}.
\]

Average spending at age \( j \) in 1950 is then computed by

\[
m_{j,1950} = \frac{m_{j,1963}}{(1 + \gamma_j)^{13}}.
\]

The results, labelled as Estimate 1, are reported in Table A1. There are at least two problems with this approach. First, average spending for the elderly (over 65 years of age) is likely to be growing at a higher rate in 1963-1977 than in 1950-1963, due to the implementation of Medicare in 1966. Second, the estimated spending are not increasing in age. In particular, there is a significant decrease in medical spending at ages 55-64 which is difficult to justify. As a partial remedy for these problems, an additional assumption is imposed to obtain a second set of estimates. The additional assumption is that average spending for the elderly had been growing at the same rate as that of the 45-54 age group during the period
1950-1963. This implies

$$m_{j,1950} = \begin{cases} \frac{m_{j,1963}}{(1+\gamma_j)^{13}} & \text{for } j < 45 \\ \frac{m_{j,1963}}{(1+\gamma_{45})^{13}} & \text{for } j \geq 45. \end{cases}$$

The resulting estimates are almost constant during ages 45-64 and increases significantly afterwards. Since the second set of estimates exhibits a more reasonable life-cycle pattern, it is used in this paper.

The age-specific medical spending is then weighted using census estimates on the size of the corresponding age group. The results for various subgroups in 1950 and 2000 are reported in Table A2. Under the second approach, the estimated level of real per-capita spending for the entire population ($432) is closer to that obtained from the aggregate statistics ($448). Also, based on Estimate 2, the growth rate of medical spending among the elderly and the non-elderly are 5.31% and 3.73% respectively.

1.7.2 The Medicare Program

The program is made up of two main parts. Part A, known as the Hospital Insurance program, paid for hospital services. It is universal and mandatory, meaning every person of age 65 or above is automatically enrolled into the program. No premium is required. Its major source of funding has been a payroll tax of 2.9%, which is split between employers and employees. Part B of the program, referred to as the Supplemental Medical Insurance program, provides benefits for physician, out-patient, emergency room and other medical services. Enrollment

\footnote{In 1997, a third part of Medicare, known as Medicare Advantage Program, was implemented. The objective of this program is to expand the beneficiaries’ choices for the supplier of health care.}
is voluntary and a monthly premium is required. The intentionally low premium had encouraged over 90% of the enrollees to participate but contributed only a small portion of the revenues. Over the years, the government has subsidized approximately 75% of the total revenues of Part B. The cost sharing structure of the Medicare program is illustrated in Table A5. For the Supplemental Medical Insurance program, all enrollees are subjected to an annual deductible and a coinsurance rate of 20%. As for the Hospital Insurance program, Medicare covers all inpatient hospital expenses for the first 60 days after an annual deductible. For days 61-90, the enrollees only have to pay 25% of the expenses. After day 90, each enrollee can draw on an additional 60-day non-renewable lifetime reserve. During these 60 days, the program covers 50% of the hospital expenses.

1.7.3 Proof of Lemma

All agents born at time $t \geq 0$ begin their lives with an empty past history of shocks, zero wealth and initial health status $\overline{h}$. The decision rule for consumption at age 0 is thus $c_0(\overline{h}, 0, z^{-1}, \varepsilon^0) = c_0(z^0)$, where $z^0 = \varepsilon^0$. The same is true for the other decision rules.

Suppose the lemma is true for age $j = 0, 1, ..., n$, with $n < J$. This means savings and post-treatment health at age $n$ are given by $s_n(z^n)$ and

$$h_{n+1}(z^n) = i \left[ m_n(z^n) \right] + (1 - \delta_h) h_n(z^{n-1}) + \varepsilon^n,$$

respectively. Then the decision rule for consumption at age $(n + 1)$ is

$$c_{n+1}[h_{n+1}(z^n), s_n(z^n), z^n, \varepsilon^{n+1}] = c_{n+1}(z^{n+1}).$$
The same argument can be applied on the initial old agents. This follows from the fact that they have empty past history of shocks, zero wealth and initial health status $\bar{h}$ at time 0. This completes the proof.
1.8 Figures and Tables

Figure 1.1: Real Per-capita Personal Medical Expenditures, 1930-2001.

Figure 1.2: Share of Medical Spending in GDP, 1930-2001.
Figure 1.3: Components of Personal Medical Expenditures, 1930-2001.

Figure 1.4: Life Expectancy at birth and Age-adjusted Mortality Rate, 1900-2001.
Figure 1.5: Survival Probability at Age 25 and above, 1949-51 and 2001.

Figure 1.6: Distribution of Personal Medical Expenditures by Source of Funds, 1950-2001.
Figure 1.7: Percentage Distribution of Deaths by Age, 1900-2001.

Figure 1.8: Age-Adjusted Mortality Rates (per 100,000 population) of Diseases of Heart and Cancers, 1900-2001.
Figure 1.9: Average Annual Incidence Rate of Coronary Heart Disease, 1987-2000.

Figure 1.10: Average Annual Incidence Rate of Cancers, 1989-2001.
Figure 1.11: Life Expectancy by Age, 1950.

Figure 1.12: Life Expectancy by Age, 2001.
Figure 1.13: Survival Probabilities in 1950.

Figure 1.14: Survival Probabilities in 2001.
Figure 1.15: Ratio of Medical Spending to Income by Age, 1950.

Figure 1.16: Ratio of Medical Spending to Income by Age, 2001.
Table 1.1
Age-specific Death Rates [per 100,000 population for specified group].

<table>
<thead>
<tr>
<th>Group</th>
<th>1900</th>
<th>1930</th>
<th>1950</th>
<th>1970</th>
<th>2001</th>
</tr>
</thead>
<tbody>
<tr>
<td>Under 1 year</td>
<td>16245</td>
<td>6900</td>
<td>3299</td>
<td>2142</td>
<td>683</td>
</tr>
<tr>
<td>1-4 years</td>
<td>1984</td>
<td>564</td>
<td>139</td>
<td>84</td>
<td>33</td>
</tr>
<tr>
<td>5-14 years</td>
<td>386</td>
<td>172</td>
<td>60</td>
<td>41</td>
<td>17</td>
</tr>
<tr>
<td>15-34 years</td>
<td>699</td>
<td>395</td>
<td>154</td>
<td>140</td>
<td>93</td>
</tr>
<tr>
<td>35-54 years</td>
<td>1216</td>
<td>913</td>
<td>580</td>
<td>523</td>
<td>308</td>
</tr>
<tr>
<td>55-64 years</td>
<td>2724</td>
<td>2403</td>
<td>1912</td>
<td>1659</td>
<td>964</td>
</tr>
<tr>
<td>over 65 years</td>
<td>8226</td>
<td>7372</td>
<td>6232</td>
<td>5892</td>
<td>5087</td>
</tr>
</tbody>
</table>


Table 1.2
Death Rates* [per 100,000 population] for Selected Infectious Diseases: 1900-1980.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Pneumonia and Influenza</td>
<td>202.2</td>
<td>155.9</td>
<td>207.8</td>
<td>102.5</td>
<td>70.3</td>
<td>31.3</td>
<td>37.3</td>
<td>30.9</td>
<td>24.1</td>
</tr>
<tr>
<td>Tuberculosis</td>
<td>194.4</td>
<td>153.8</td>
<td>113.1</td>
<td>71.1</td>
<td>45.9</td>
<td>22.5</td>
<td>6.1</td>
<td>2.6</td>
<td>0.9</td>
</tr>
<tr>
<td>Typhoid Fever</td>
<td>31.3</td>
<td>22.5</td>
<td>7.6</td>
<td>4.8</td>
<td>1.1</td>
<td>0.1</td>
<td>**</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>Diphtheria</td>
<td>40.3</td>
<td>21.1</td>
<td>15.3</td>
<td>4.9</td>
<td>1.1</td>
<td>0.3</td>
<td>**</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>Whooping Cough</td>
<td>12.2</td>
<td>11.6</td>
<td>12.5</td>
<td>4.8</td>
<td>2.2</td>
<td>0.7</td>
<td>0.1</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>Measles</td>
<td>13.3</td>
<td>12.4</td>
<td>8.8</td>
<td>3.2</td>
<td>0.5</td>
<td>0.3</td>
<td>0.2</td>
<td>**</td>
<td>**</td>
</tr>
</tbody>
</table>

* The reported data are not age-adjusted. This is because age-adjusted mortality rates for some diseases are not available for early years.

** means less than 0.05.
Table 1.3a
Baseline Parameterization.

<table>
<thead>
<tr>
<th>Demography</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum age</td>
<td>$J$</td>
<td>15</td>
</tr>
<tr>
<td>Retirement age</td>
<td>$I$</td>
<td>7</td>
</tr>
<tr>
<td>Parameters in survival probability function</td>
<td>$\psi$</td>
<td>0.00112*</td>
</tr>
<tr>
<td></td>
<td>$\theta$</td>
<td>1.892*</td>
</tr>
<tr>
<td>Health endowment</td>
<td>$\bar{h}$</td>
<td>100</td>
</tr>
<tr>
<td>Population growth rate (5-year)</td>
<td>$\gamma$</td>
<td>6.4%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Preferences</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient of relative risk aversion</td>
<td>$\sigma$</td>
<td>0.97</td>
</tr>
<tr>
<td>Subjective discount factor (5-year)</td>
<td>$\beta$</td>
<td>0.904</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Production of Goods</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Capital’s share of income</td>
<td>$\alpha$</td>
<td>0.33</td>
</tr>
<tr>
<td>Depreciation rate of physical capital (5-year)</td>
<td>$\delta_k$</td>
<td>0.41</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Production of Health</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative price of medical care, 1950</td>
<td>$p_{1950}$</td>
<td>10.0</td>
</tr>
<tr>
<td>Relative price of medical care, 2001</td>
<td>$p_{2001}$</td>
<td>28.8</td>
</tr>
<tr>
<td>Productivity of medical care, 1950</td>
<td>$\eta_{1950}$</td>
<td>0.274*</td>
</tr>
<tr>
<td>Productivity of medical care, 2001</td>
<td>$\eta_{2001}$</td>
<td>1.510</td>
</tr>
<tr>
<td>Parameter in health production function</td>
<td>$\xi$</td>
<td>0.10*</td>
</tr>
<tr>
<td>Depreciation of health</td>
<td>$\delta_h$</td>
<td>0.065*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Morbidity</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Health shock in “healthy” state</td>
<td>$\varepsilon_1$</td>
<td>0</td>
</tr>
<tr>
<td>Health shock in “sick” state</td>
<td>$\varepsilon_2$</td>
<td>$-20^*$</td>
</tr>
</tbody>
</table>

*The values are obtained from the minimization process.
Table 1.3b
Baseline Parameterization: Labor Endowment.

<table>
<thead>
<tr>
<th>Age</th>
<th>25-29</th>
<th>30-34</th>
<th>35-39</th>
<th>40-44</th>
<th>45-49</th>
<th>50-54</th>
<th>55-59</th>
<th>60-64</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year 1950</td>
<td>1.00</td>
<td>1.16</td>
<td>1.21</td>
<td>1.22</td>
<td>1.23</td>
<td>1.24</td>
<td>1.20</td>
<td>1.11</td>
</tr>
<tr>
<td>Year 2001</td>
<td>2.43</td>
<td>2.98</td>
<td>3.25</td>
<td>3.42</td>
<td>3.54</td>
<td>3.47</td>
<td>3.66</td>
<td>3.37</td>
</tr>
</tbody>
</table>

Table 1.3c
Baseline Parameterization: Probability of Health States.

<table>
<thead>
<tr>
<th>Age</th>
<th>Healthy</th>
<th>Chronic Diseases</th>
</tr>
</thead>
<tbody>
<tr>
<td>25-29</td>
<td>0.997</td>
<td>0.003</td>
</tr>
<tr>
<td>30-34</td>
<td>0.996</td>
<td>0.004</td>
</tr>
<tr>
<td>35-39</td>
<td>0.991</td>
<td>0.009</td>
</tr>
<tr>
<td>40-44</td>
<td>0.984</td>
<td>0.016</td>
</tr>
<tr>
<td>45-49</td>
<td>0.974</td>
<td>0.026</td>
</tr>
<tr>
<td>50-54</td>
<td>0.958</td>
<td>0.042</td>
</tr>
<tr>
<td>55-59</td>
<td>0.936</td>
<td>0.064</td>
</tr>
<tr>
<td>60-64</td>
<td>0.909</td>
<td>0.091</td>
</tr>
<tr>
<td>65-69</td>
<td>0.876</td>
<td>0.124</td>
</tr>
<tr>
<td>70-74</td>
<td>0.845</td>
<td>0.155</td>
</tr>
<tr>
<td>75-79</td>
<td>0.821</td>
<td>0.179</td>
</tr>
<tr>
<td>80-84</td>
<td>0.801</td>
<td>0.199</td>
</tr>
<tr>
<td>85-89</td>
<td>0.789</td>
<td>0.211</td>
</tr>
<tr>
<td>90-94</td>
<td>0.763</td>
<td>0.237</td>
</tr>
<tr>
<td>95-99</td>
<td>0.732</td>
<td>0.268</td>
</tr>
</tbody>
</table>
Table 1.4
Selected Real-World Statistics.

<table>
<thead>
<tr>
<th></th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Share of Medical Spending in GDP, 1950</td>
<td>3.7%</td>
</tr>
<tr>
<td>2</td>
<td>Ratio of Medical Spending to Income Among the Non-Elderly, 1950</td>
<td>0.037</td>
</tr>
<tr>
<td>3</td>
<td>Ratio of Medical Spending to Income Among the Elderly, 1950</td>
<td>0.117</td>
</tr>
<tr>
<td>4</td>
<td>% of Out-of-Pocket Expenditures in Total Medical Spending, 1950</td>
<td>68.3%</td>
</tr>
<tr>
<td>5</td>
<td>% of Out-of-Pocket Expenditures in Total Medical Spending, 2001</td>
<td>16.6%</td>
</tr>
<tr>
<td>6</td>
<td>Tax rate used to finance Medicare in 2001</td>
<td>2.9%</td>
</tr>
<tr>
<td>7</td>
<td>Life Expectancy at age 25 in 1950</td>
<td>47.7</td>
</tr>
<tr>
<td>8</td>
<td>Life Expectancy at age 45 in 1950</td>
<td>29.1</td>
</tr>
<tr>
<td>9</td>
<td>Life Expectancy at age 65 in 1950</td>
<td>14.3</td>
</tr>
</tbody>
</table>
Table 1.5
Main Findings from the Benchmark Model.

<table>
<thead>
<tr>
<th></th>
<th>Data 1950</th>
<th>Data 2001</th>
<th>Model 1950</th>
<th>Model 2001</th>
</tr>
</thead>
<tbody>
<tr>
<td>Share of Medical Spending in GDP</td>
<td>3.7% 12.2%</td>
<td>3.7% 12.4%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Share of Medical Spending in PCE</td>
<td>5.6% 17.3%</td>
<td>4.8% 16.6%</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Real Per-capita Medical Spending</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Annual Growth Rate</td>
<td>4.44%*</td>
<td>4.43%</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Medical Spending by Source of Payment</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Out of Pocket</td>
<td>68.3% 16.6%</td>
<td>68.2% 24.7%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicare</td>
<td>0% 19.4%</td>
<td>0% 15.5%</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Life Expectancy by Age</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age 25</td>
<td>47.7 54.2</td>
<td>47.7 51.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age 45</td>
<td>29.1 35.0</td>
<td>28.9 32.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age 65</td>
<td>14.3 18.2</td>
<td>14.7 17.8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* The reported value is the growth rate of real per-capita spending among those aged 25 or above during the period 1950-2000. See Footnote 3 for details.
Table 1.6
Changes in Average Life Expectancy between 1950 and 2001.\textsuperscript{a,b}

<table>
<thead>
<tr>
<th></th>
<th>Data</th>
<th>Model</th>
<th>% Explained by the Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measure 1</td>
<td>3.35</td>
<td>2.69</td>
<td>80.4%</td>
</tr>
<tr>
<td>Measure 2</td>
<td>5.58</td>
<td>3.72</td>
<td>66.7%</td>
</tr>
<tr>
<td>Measure 3</td>
<td>5.35</td>
<td>3.65</td>
<td>68.1%</td>
</tr>
</tbody>
</table>

\textsuperscript{a}The reported values are in number of years.
\textsuperscript{b}See footnote 30 for the definitions.

Table 1.7
Findings: Life-cycle Pattern of Medical Spending.
A. Ratio of Medical Spending to Income by Age.*

<table>
<thead>
<tr>
<th></th>
<th>1950 Data</th>
<th>1950 Model</th>
<th>2001 Data</th>
<th>2001 Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>25-34</td>
<td>0.035</td>
<td>0.033</td>
<td>0.077</td>
<td>0.111</td>
</tr>
<tr>
<td>35-44</td>
<td>0.033</td>
<td>0.041</td>
<td>0.074</td>
<td>0.124</td>
</tr>
<tr>
<td>45-54</td>
<td>0.040</td>
<td>0.046</td>
<td>0.098</td>
<td>0.134</td>
</tr>
<tr>
<td>55-64</td>
<td>0.046</td>
<td>0.043</td>
<td>0.161</td>
<td>0.204</td>
</tr>
<tr>
<td>\textit{Under 65}</td>
<td>0.037</td>
<td>0.040</td>
<td>0.096</td>
<td>0.123</td>
</tr>
<tr>
<td>\textit{Over 65}</td>
<td>0.117</td>
<td>0.105</td>
<td>0.612</td>
<td>0.635</td>
</tr>
</tbody>
</table>

B. Average Annual Growth Rate of Medical Spending by Age.

<table>
<thead>
<tr>
<th></th>
<th>Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>25-64</td>
<td>3.81%</td>
<td>4.26%</td>
</tr>
<tr>
<td>65-74</td>
<td>5.18%</td>
<td>5.09%</td>
</tr>
<tr>
<td>75+</td>
<td>5.28%</td>
<td>6.49%</td>
</tr>
<tr>
<td>\textit{Over 65}</td>
<td>5.41%</td>
<td>5.48%</td>
</tr>
<tr>
<td>\textit{Over 25}</td>
<td>4.44%</td>
<td>4.43%</td>
</tr>
</tbody>
</table>

* See Footnote 29 for details.
** The reported values are for the year 2000.
Table 1.8
Benchmark Economy vs Economy with No Insurance.

<table>
<thead>
<tr>
<th>Medical Spending</th>
<th>1950</th>
<th>2001</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No Insurance</td>
<td>Benchmark</td>
</tr>
<tr>
<td>% in GDP</td>
<td>2.1%</td>
<td>3.7%</td>
</tr>
<tr>
<td>% in PCE</td>
<td>2.7%</td>
<td>4.8%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Life Expectancy</th>
<th>1950</th>
<th>2001</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age 25</td>
<td>47.5</td>
<td>47.7</td>
</tr>
<tr>
<td>Age 45</td>
<td>28.8</td>
<td>28.9</td>
</tr>
<tr>
<td>Age 65</td>
<td>14.6</td>
<td>14.7</td>
</tr>
</tbody>
</table>

Table 1.9a
Effects of Changing the Coefficient of Relative Risk Aversion.

<table>
<thead>
<tr>
<th>Coefficient of Relative Risk Aversion</th>
<th>0.95</th>
<th>0.96</th>
<th>0.97</th>
<th>0.98</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benchmark Economy, 1950</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Share of Medical Spending in GDP</td>
<td>2.1%</td>
<td>2.5%</td>
<td><strong>3.7%</strong></td>
<td>5.5%</td>
</tr>
<tr>
<td>Share of Medical Spending in PCE</td>
<td>2.7%</td>
<td>3.3%</td>
<td><strong>4.8%</strong></td>
<td>7.1%</td>
</tr>
<tr>
<td>Units of Insurance Purchased</td>
<td>0.016</td>
<td>0.020</td>
<td><strong>0.028</strong></td>
<td>0.042</td>
</tr>
<tr>
<td>Benchmark Economy, 2001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Share of Medical Spending in GDP</td>
<td><strong>8.6%</strong></td>
<td>10.1%</td>
<td><strong>12.4%</strong></td>
<td>16.9%</td>
</tr>
<tr>
<td>Share of Medical Spending in PCE</td>
<td>11.6%</td>
<td>13.6%</td>
<td><strong>16.6%</strong></td>
<td>22.5%</td>
</tr>
<tr>
<td>Units of Insurance Purchased</td>
<td>0.209</td>
<td>0.254</td>
<td><strong>0.320</strong></td>
<td>0.458</td>
</tr>
</tbody>
</table>

Note: The numbers in bold are the baseline results.
Table 1.9b
Effects of Changing the Coefficient of Relative Risk Aversion.

<table>
<thead>
<tr>
<th>Coefficient of Relative Risk Aversion</th>
<th>0.95</th>
<th>0.96</th>
<th><strong>0.97</strong></th>
<th>0.98</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Benchmark Economy, 1950</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Productivity of Medical Care, $\eta_{1950}$</td>
<td>0.470</td>
<td>0.380</td>
<td><strong>0.274</strong></td>
<td>0.200</td>
</tr>
<tr>
<td>Share of Medical Spending in GDP</td>
<td>3.7%</td>
<td>3.7%</td>
<td><strong>3.7%</strong></td>
<td>3.7%</td>
</tr>
<tr>
<td>Share of Medical Spending in PCE</td>
<td>4.8%</td>
<td>4.9%</td>
<td><strong>4.8%</strong></td>
<td>4.9%</td>
</tr>
<tr>
<td><strong>Benchmark Economy, 2001</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Productivity of Medical Care, $\eta_{2001}$</td>
<td>2.59</td>
<td>2.09</td>
<td><strong>1.51</strong></td>
<td>1.10</td>
</tr>
<tr>
<td>Share of Medical Spending in GDP</td>
<td>11.8%</td>
<td>12.3%</td>
<td><strong>12.4%</strong></td>
<td>13.6%</td>
</tr>
<tr>
<td>Share of Medical Spending in PCE</td>
<td>15.8%</td>
<td>16.4%</td>
<td><strong>16.6%</strong></td>
<td>18.0%</td>
</tr>
</tbody>
</table>

Note: The numbers in bold are the baseline results.
*The productivity of medical care in 2001 is given by $\eta_{2001} = (1.034)^{51} \eta_{1950}$.

Table 1.10
Effects of Changing the Price of Medical Care in 1950 ($p_{1950}$)

<table>
<thead>
<tr>
<th>Price of Medical Care in 1950</th>
<th>5</th>
<th>10</th>
<th>15</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Benchmark Economy, 1950</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Share of Medical Spending in GDP</td>
<td>3.9%</td>
<td><strong>3.7%</strong></td>
<td>3.5%</td>
<td>3.4%</td>
</tr>
<tr>
<td>Share of Medical Spending in PCE</td>
<td>5.2%</td>
<td><strong>4.8%</strong></td>
<td>4.6%</td>
<td>4.4%</td>
</tr>
</tbody>
</table>

| **Benchmark Economy, 2001**  |   |    |    |    |
| Share of Medical Spending in GDP | 13.0% | **12.4%** | 12.1% | 11.9% |
| Share of Medical Spending in PCE | 17.3% | **16.6%** | 16.2% | 15.9% |

Note: The numbers in bold are the baseline results.
*The price of medical care in 2001 is given by $p_{2001} = (1.021)^{51} p_{1950}$.
Table A1
Real Per-Person Medical Spending by Age, 1950-2000.\textsuperscript{a}

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0 to 4 years</td>
<td>104</td>
<td>104</td>
<td>241</td>
<td>550</td>
<td>594</td>
<td>1176</td>
<td>1636</td>
<td>1500</td>
</tr>
<tr>
<td>5 to 14 years</td>
<td>133</td>
<td>133</td>
<td>240</td>
<td>336</td>
<td>455</td>
<td>690</td>
<td>1005</td>
<td>1446</td>
</tr>
<tr>
<td>15 to 24 years</td>
<td>397</td>
<td>397</td>
<td>574</td>
<td>1046</td>
<td>856</td>
<td>1193</td>
<td>2361</td>
<td>1651</td>
</tr>
<tr>
<td>25 to 34 years</td>
<td>502</td>
<td>502</td>
<td>744</td>
<td>932</td>
<td>1139</td>
<td>1382</td>
<td>1746</td>
<td>2351</td>
</tr>
<tr>
<td>35 to 44 years</td>
<td>507</td>
<td>507</td>
<td>763</td>
<td>768</td>
<td>1184</td>
<td>1441</td>
<td>2253</td>
<td>2847</td>
</tr>
<tr>
<td>45 to 54 years</td>
<td>589</td>
<td>589</td>
<td>942</td>
<td>1087</td>
<td>1562</td>
<td>2309</td>
<td>3042</td>
<td>4045</td>
</tr>
<tr>
<td>55 to 64 years</td>
<td>419</td>
<td>590</td>
<td>944</td>
<td>1740</td>
<td>2264</td>
<td>3195</td>
<td>5072</td>
<td>5736</td>
</tr>
<tr>
<td>65 to 74 years</td>
<td>434</td>
<td>715</td>
<td>1143</td>
<td>2229</td>
<td>3245</td>
<td>5957</td>
<td>7105</td>
<td>8950</td>
</tr>
<tr>
<td>75+ years</td>
<td>563</td>
<td>1184</td>
<td>1895</td>
<td>4309</td>
<td>6996</td>
<td>11325</td>
<td>16253</td>
<td>15503</td>
</tr>
</tbody>
</table>

\textsuperscript{a}The reported values are in constant 2001 dollars.
# Table A2
Real Per-Person Medical Spending for Selected Age Groups, 1950 and 2000.

<table>
<thead>
<tr>
<th>Age Group</th>
<th>1950 Estimate 1</th>
<th>1950 Estimate 2</th>
<th>2000</th>
<th>Average Annual Growth Rate**</th>
</tr>
</thead>
<tbody>
<tr>
<td>25-64 years</td>
<td>509</td>
<td>539</td>
<td>3,492</td>
<td>3.81%</td>
</tr>
<tr>
<td>over 65 years</td>
<td>475</td>
<td>863</td>
<td>12,053</td>
<td>5.41%</td>
</tr>
<tr>
<td>over 25 years</td>
<td>504</td>
<td>584</td>
<td>5,140</td>
<td>4.44%</td>
</tr>
<tr>
<td>All Ages</td>
<td>385</td>
<td>432</td>
<td>3,868</td>
<td>4.48%</td>
</tr>
<tr>
<td></td>
<td>(448)</td>
<td>(448)</td>
<td>(4,361)</td>
<td>(4.66%)</td>
</tr>
</tbody>
</table>

*Elderly’s Share of Medical Spending*

<table>
<thead>
<tr>
<th>Age Group</th>
<th>1950</th>
<th>2000</th>
<th>Average Annual Growth Rate**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medical Spending*</td>
<td>13.1%</td>
<td>20.6%</td>
<td>45.1%</td>
</tr>
<tr>
<td>Population*</td>
<td>13.9%</td>
<td>13.9%</td>
<td>19.3%</td>
</tr>
</tbody>
</table>

Note: The reported values are in constant 2001 dollars. The values in parentheses are per-person values obtained by dividing total personal medical expenditures by the total civilian population.

* These are the elderly’s share among those aged 25 and above.
** Average annual growth rate over the period 1950-2001 based on Estimate 2.
Table A3
Age-specific Death Rates [per 100,000 population for each group], 1950.

<table>
<thead>
<tr>
<th>Age Group</th>
<th>All Causes</th>
<th>Accident</th>
<th>Suicide</th>
<th>Homicide</th>
<th>Subtotal(^a)</th>
<th>% Share(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Ages</td>
<td>936.8</td>
<td>60.6</td>
<td>11.4</td>
<td>5.3</td>
<td>77.3</td>
<td>8.3%</td>
</tr>
<tr>
<td>Under 1</td>
<td>3299.2</td>
<td>114.2</td>
<td>0.0</td>
<td>4.4</td>
<td>118.6</td>
<td>3.6%</td>
</tr>
<tr>
<td>1 to 4</td>
<td>139.4</td>
<td>36.8</td>
<td>0.0</td>
<td>0.6</td>
<td>37.4</td>
<td>26.8%</td>
</tr>
<tr>
<td>5 to 14</td>
<td>60.1</td>
<td>22.7</td>
<td>0.2</td>
<td>0.5</td>
<td>23.4</td>
<td>38.9%</td>
</tr>
<tr>
<td>15 to 24</td>
<td>128.1</td>
<td>54.8</td>
<td>4.5</td>
<td>6.3</td>
<td>65.6</td>
<td>51.2%</td>
</tr>
<tr>
<td>25 to 34</td>
<td>178.7</td>
<td>45.7</td>
<td>9.1</td>
<td>9.9</td>
<td>64.7</td>
<td>36.2%</td>
</tr>
<tr>
<td>35 to 44</td>
<td>358.7</td>
<td>45.7</td>
<td>14.3</td>
<td>8.8</td>
<td>68.8</td>
<td>19.2%</td>
</tr>
<tr>
<td>45 to 54</td>
<td>853.9</td>
<td>53.0</td>
<td>20.9</td>
<td>6.1</td>
<td>80.0</td>
<td>9.4%</td>
</tr>
<tr>
<td>55 to 64</td>
<td>1901.0</td>
<td>70.8</td>
<td>26.8</td>
<td>4.0</td>
<td>101.6</td>
<td>5.3%</td>
</tr>
<tr>
<td>65 to 74</td>
<td>4104.3</td>
<td>116.9</td>
<td>29.6</td>
<td>3.2</td>
<td>149.7</td>
<td>3.6%</td>
</tr>
<tr>
<td>75 to 84</td>
<td>9331.1</td>
<td>315.7</td>
<td>31.1</td>
<td>2.6</td>
<td>349.4</td>
<td>3.7%</td>
</tr>
<tr>
<td>85 above</td>
<td>20196.9</td>
<td>972.6</td>
<td>28.8</td>
<td>0.0</td>
<td>1001.4</td>
<td>5.0%</td>
</tr>
</tbody>
</table>


\(^a\) This is the sum of all deaths due to accident, suicide and homicide.

\(^b\) This is the share of deaths due to accident, suicide and homicide in deaths of all causes.
Table A4
Age-specific Death Rates [per 100,000 population for each group], 2001.

<table>
<thead>
<tr>
<th>Age Group</th>
<th>All Causes</th>
<th>Accident</th>
<th>Suicide</th>
<th>Homicide</th>
<th>Subtotal&lt;sup&gt;a&lt;/sup&gt;</th>
<th>% Share&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Ages</td>
<td>848.5</td>
<td>35.7</td>
<td>10.8</td>
<td>7.1</td>
<td>53.6</td>
<td>6.3%</td>
</tr>
<tr>
<td>Under 1</td>
<td>683.4</td>
<td>24.2</td>
<td>0.0</td>
<td>8.2</td>
<td>32.4</td>
<td>4.7%</td>
</tr>
<tr>
<td>1 to 4</td>
<td>33.3</td>
<td>11.2</td>
<td>0.0</td>
<td>2.7</td>
<td>13.9</td>
<td>41.7%</td>
</tr>
<tr>
<td>5 to 9</td>
<td>15.3</td>
<td>6.4</td>
<td>0.0</td>
<td>0.7</td>
<td>7.1</td>
<td>46.4%</td>
</tr>
<tr>
<td>10 to 14</td>
<td>19.2</td>
<td>7.4</td>
<td>1.3</td>
<td>0.9</td>
<td>9.6</td>
<td>50.0%</td>
</tr>
<tr>
<td>15 to 19</td>
<td>66.9</td>
<td>32.8</td>
<td>7.9</td>
<td>9.4</td>
<td>50.1</td>
<td>74.9%</td>
</tr>
<tr>
<td>20 to 24</td>
<td>95.0</td>
<td>39.5</td>
<td>12.0</td>
<td>17.3</td>
<td>68.8</td>
<td>72.4%</td>
</tr>
<tr>
<td>25 to 29</td>
<td>96.2</td>
<td>31.4</td>
<td>12.6</td>
<td>14.9</td>
<td>58.9</td>
<td>61.2%</td>
</tr>
<tr>
<td>30 to 34</td>
<td>113.5</td>
<td>28.5</td>
<td>13.0</td>
<td>11.5</td>
<td>53.0</td>
<td>46.7%</td>
</tr>
<tr>
<td>35 to 39</td>
<td>165.9</td>
<td>34.1</td>
<td>14.3</td>
<td>10.4</td>
<td>58.8</td>
<td>35.4%</td>
</tr>
<tr>
<td>40 to 44</td>
<td>240.5</td>
<td>36.7</td>
<td>15.7</td>
<td>7.0</td>
<td>58.7</td>
<td>16.5%</td>
</tr>
<tr>
<td>45 to 49</td>
<td>354.8</td>
<td>36.0</td>
<td>15.7</td>
<td>7.0</td>
<td>58.7</td>
<td>16.5%</td>
</tr>
<tr>
<td>50 to 54</td>
<td>512.4</td>
<td>31.9</td>
<td>14.6</td>
<td>5.5</td>
<td>52.0</td>
<td>10.1%</td>
</tr>
<tr>
<td>55 to 59</td>
<td>771.8</td>
<td>29.3</td>
<td>14.0</td>
<td>4.4</td>
<td>47.7</td>
<td>6.2%</td>
</tr>
<tr>
<td>60 to 64</td>
<td>1,210.7</td>
<td>31.5</td>
<td>12.0</td>
<td>3.5</td>
<td>47.0</td>
<td>3.9%</td>
</tr>
<tr>
<td>65 to 69</td>
<td>1,869.7</td>
<td>36.8</td>
<td>12.7</td>
<td>3.3</td>
<td>52.8</td>
<td>2.8%</td>
</tr>
<tr>
<td>70 to 74</td>
<td>2,878.3</td>
<td>49.3</td>
<td>13.9</td>
<td>2.5</td>
<td>65.7</td>
<td>2.3%</td>
</tr>
<tr>
<td>75 to 79</td>
<td>4,494.0</td>
<td>80.4</td>
<td>16.4</td>
<td>2.5</td>
<td>99.3</td>
<td>2.2%</td>
</tr>
<tr>
<td>80 to 84</td>
<td>7,151.9</td>
<td>130.4</td>
<td>18.9</td>
<td>2.5</td>
<td>151.8</td>
<td>2.1%</td>
</tr>
<tr>
<td>85 above</td>
<td>15,112.8</td>
<td>276.4</td>
<td>17.5</td>
<td>2.4</td>
<td>296.3</td>
<td>2.0%</td>
</tr>
</tbody>
</table>


<sup>a</sup> This is the sum of all deaths due to accident, suicide and homicide.

<sup>b</sup> This is the share of deaths due to accident, suicide and homicide in deaths of all causes.
Table A5

<table>
<thead>
<tr>
<th>Year</th>
<th>Hospital Insurance</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Day 1-60</td>
<td>Day 61-90</td>
<td>Day 91-150</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(deductible only)</td>
<td>(coinsurance)</td>
<td>(lifetime reserve)</td>
<td></td>
</tr>
<tr>
<td>1966</td>
<td>$40</td>
<td>75%</td>
<td>50%</td>
<td></td>
</tr>
<tr>
<td>1970</td>
<td>$52</td>
<td>75%</td>
<td>50%</td>
<td></td>
</tr>
<tr>
<td>1980</td>
<td>$180</td>
<td>75%</td>
<td>50%</td>
<td></td>
</tr>
<tr>
<td>1985</td>
<td>$400</td>
<td>75%</td>
<td>50%</td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td>$592</td>
<td>75%</td>
<td>50%</td>
<td></td>
</tr>
<tr>
<td>1995</td>
<td>$715</td>
<td>75%</td>
<td>50%</td>
<td></td>
</tr>
<tr>
<td>2001</td>
<td>$792</td>
<td>75%</td>
<td>50%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>Supplemental Medical Insurance</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Annual Deductible</td>
<td>Coinsurance Rate</td>
<td>Monthly Premium</td>
<td></td>
</tr>
<tr>
<td>1966</td>
<td>$50</td>
<td>80%</td>
<td>$3.00</td>
<td></td>
</tr>
<tr>
<td>1970</td>
<td>$50</td>
<td>80%</td>
<td>$4.00</td>
<td></td>
</tr>
<tr>
<td>1980</td>
<td>$60</td>
<td>80%</td>
<td>$8.70</td>
<td></td>
</tr>
<tr>
<td>1985</td>
<td>$75</td>
<td>80%</td>
<td>$15.50</td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td>$75</td>
<td>80%</td>
<td>$28.60</td>
<td></td>
</tr>
<tr>
<td>1995</td>
<td>$100</td>
<td>80%</td>
<td>$46.10</td>
<td></td>
</tr>
<tr>
<td>2001</td>
<td>$100</td>
<td>80%</td>
<td>$50.00</td>
<td></td>
</tr>
</tbody>
</table>
Chapter 2
The Welfare Implications of Increased Medical Spending

2.1 Introduction

Over the period 1950-2001, the share of personal medical expenditures in GDP increased from 3.7% to 12.2%. This tremendous increase in medical spending was accompanied by a remarkable expansion in insurance coverage. In 1950, 68.3% of total personal medical expenditures was paid directly by the consumers. This decreased to 17% in 2001. Over the same period, the share paid by private insurance increased from 8.5% to 35%. Previous work in the literature argued that, due to moral hazard problems, the prevalence of health insurance is a major driving force behind the rapid growth in medical spending.1 Moral hazard occurs because it is difficult to measure the extent of health loss due to a particular episode of illness. This means it is not practicable to make reimbursement dependent on the severity of illness. Conventionally, the level of expenses incurred during medical treatment is used as a reflection on how severe an illness is, and reimbursement

1See, for instance, Pauly (1986) for a review of this literature.
is paid according to these expenses. An incentive problem arises because medical spending is to some extent controlled by the insured. When health insurance reimburses part of the medical spending, it effectively lowers the marginal cost of medical care. The purchase of insurance thus induces people to spend more than they would choose to if they were uninsured. The situation is exacerbated by the fact that employer-provided health insurance is tax-deductible, which leads to distortions in the demand for health insurance. Thus, according to the previous literature, the tremendous increase in medical expenditures is inefficient and represents a significant welfare loss.

The objective of this chapter is to re-examine this welfare issue. This is achieved in two steps. The first step is to construct a benchmark for welfare analysis. The benchmark economy is a complete-market version of the model presented in section 1.3. Recall that in that model, consumers face two types of uncertainty: uncertainty concerning their health status and uncertainty with respect to the length of their lifetimes. Markets are complete in the sense that there are enough assets via which consumers can perfectly insure against any fluctuation in consumption brought about by these uncertainties. These assets include a complete set of state-contingent insurance contracts and an actuarially fair annuity. A key underlying assumption is that information is perfect, so everyone’s health status is observable at all times. Clearly, there is no incentive problem under this assumption. In the theoretical analysis presented in section 2.3, it is shown that Pareto optimality can be achieved in equilibrium. The second step is to compare the welfare implications of this and the incomplete-market economy constructed in Chapter 1. Since the two are identical except for their market structures, the

---

2Moral hazard can also arise when the purchase of health insurance discourages the use of preventive medical care and thus increases the probability of getting sick. Since preventive medical care is not considered in this chapter, this sort of moral hazard is not addressed.
differences in welfare would capture the extent of welfare loss due to market incompleteness. In the quantitative analysis, the calibration procedure described in section 1.4 is applied to the complete-market economy. The calibrated version of the model yields a large increase in medical spending over time. This reinforces the conclusion reached in the previous chapter, namely that the increase in medical spending is not driven by factors associated with market structures, but by factors underlying the production and accumulation of health. In the welfare analysis presented in section 2.5, an equivalent variation measure in consumption is used to quantify the welfare difference. This measure is based on the percentage change in annual consumption required to leave a typical consumer indifferent between the two market structures. The required changes are 0.28% in 1950 and 0.76% in 2001.3

There are two major differences between this and the previous work in the literature. First, most (if not all) of the previous studies focused on partial equilibrium in a static environment. In this chapter, the welfare issue is addressed under a dynamic general equilibrium framework. Second, earlier studies estimate the extent of welfare loss indirectly through estimating the price elasticity of medical care demand.4 The rationale is that since health insurance distorts the price of medical care, a more elastic demand implies a potentially higher degree of distortion and hence a larger extent of welfare loss. In this chapter, the extent of welfare loss is measured directly using an equivalent variation measure in consumption.

The reminder of this chapter is organized as follows. Section 2.2 describes the complete-market economy. Section 2.3 reports the results obtained from the theoretical analysis, while section 2.4 reports the main findings obtained from the

---

3 Since there are other forms of market incompleteness in the model economy, these number give the upper bound on the potential welfare loss due to moral hazard problems.

4 A comprehensive survey of this literature can be found in Cutler and Zeckhauser (2000).
quantitative analysis. The welfare analysis is conducted in section 2.5. This is followed by a brief conclusion in section 2.6.

2.2 The Model

2.2.1 Market Structure

The fundamentals of this economy are described in section 1.3.1, hence they are not repeated in here. In the complete-market economy, there are a complete set of state-contingent claims, with which agents can insure against the health shock, and a perfect annuity, with which agents can insure against the mortality risk. A key underlying assumption is that health shocks experienced by the agents, as well as their health status, can be publicly observed. Consequently, both insurance and annuity contracts can be made contingent on these information.

In the contingent claims markets each claim yields one unit of current-period consumption if a particular health shock sequence occurs in the current period.\(^5\) All contingent claims are in zero net supply. Consider an age-\(j\) agent with past history of shocks, \(z_{-1}\). He will trade claims which are contingent on \((z_{-1}, \varepsilon_1), \ldots, (z_{-1}, \varepsilon_S)\).

To insure against uncertainty on the length of lifetime, an agent can save by purchasing annuity. An annuity is an asset that pays positive returns in the subsequent period if the owner survives and zero otherwise. It is offered by competitive firms in the annuity market. Under full information, the buyer’s history of illness and health status are known by these firms. This means annuity contracts with returns contingent on the buyers’ health status can be written. The firms then invest their revenues on physical capital. Following the work of Yaari (1965), the annuity is assumed to be actuarially fair. It follows that the annuity would yield a

\(^5\)Hence, unlike the usual setting, the purchase of contingent claims is an intratemporal decision rather than an intertemporal one.
higher return than physical capital and all agents would choose to fully annuitize their savings.

The sequence of events is as follows: At the beginning of each period, an agent chooses a portfolio of contingent claims and a set of state-contingent decision rules for consumption, medical care and savings. All these decisions are made before the current health shock is realized. Once it is realized, the result is publicly known. Commodities specified by the claims and the decision rules are then delivered. Before the end of the period, the agent is informed whether or not he will survive in the next period.

2.2.2 Individual’s Problem

Before the realization of the current health shock, an age-\(j\) individual is characterized by \(\theta = (h, s_{-1}, z_{-1})\), where \(h\) is the beginning-of-period health status, \(s_{-1}\) denotes previous savings in annuity and \(z_{-1}\) the past history of health shocks. Given \(h\), the gross return from annuity is \(R(h)\). The labor income of this agent is \(w_{ej}\), where \(w\) is the market wage rate for effective unit of labor. Let \(\{q_j(\varepsilon|z_{-1})|\varepsilon \in E\}\) be the prices of claims that are contingent on \((z_{-1}, \varepsilon_1), \ldots, (z_{-1}, \varepsilon_S)\) and \(\{x_j(\varepsilon|z_{-1})|\varepsilon \in E\}\) an arbitrary portfolio. The choice of contingent claims is subject to the following constraint,

\[
\sum_{\varepsilon \in E} q_j(\varepsilon|z_{-1}) x_j(\varepsilon|z_{-1}) \geq 0. \tag{2.1}
\]

The above condition requires the net value of the portfolio of claims be nonnegative. This means agents cannot borrow from the contingent claims markets. One problem with this form of borrowing is that if the borrower dies before the beginning of the next period, his loans will not be repaid. Since it is not possible to guarantee the repayment of these loans, (2.1) is imposed to rule them out.
Let $E^j$ be the expectation operator based on all information available at age $j$. The agent’s problem is to choose a portfolio of contingent claims and a set of state-contingent decision rules for consumption, medical care and savings, denoted by $\varphi = \{x_j(\varepsilon|z_{-1}), c_j(\theta, \varepsilon), m_j(\theta, \varepsilon), s_j(\theta, \varepsilon) | \varepsilon \in \mathcal{E}\}$, in order to maximize the expected value for the remaining lifetime. Formally, this is written as

$$E^j [V^j(\theta, \varepsilon)] = \max_{\varphi} E^j \left\{ U[c_j(\theta, \varepsilon)] + \beta \Phi[h_{j+1}(\theta, \varepsilon)] E^{j+1} [V^{j+1}(\theta', \varepsilon')] \right\}$$

(P2.1)

subject to

$$c_j(\theta, \varepsilon) + pm_j(\theta, \varepsilon) + s_j(\theta, \varepsilon) + \sum_{\varepsilon \in \mathcal{E}} q_j(\varepsilon|z_{-1}) x_j(\varepsilon|z_{-1}) = we_j + R(h) s_{-1} + x_j(\varepsilon|z_{-1}),$$

$$h_{j+1}(\theta, \varepsilon) = i[m_j(\theta, \varepsilon)] + (1 - \delta_h) h + \varepsilon,$$

for $\varepsilon \in \mathcal{E}$, and constraint (2.1).

The optimal condition for contingent claims is given by

$$\frac{\pi_j(\varepsilon) U'[c_j(\theta, \varepsilon)]}{q_j(\varepsilon|z_{-1})} = \frac{\pi_j(\varepsilon') U'[c_j(\theta, \varepsilon')]}{q_j(\varepsilon'|z_{-1})},$$

(2.3)

for any $\varepsilon, \varepsilon' \in \mathcal{E}$. Optimality is achieved when an additional unit of spending on all claims yields the same expected marginal utility.

The optimization problem is complicated by the fact that $\Phi(h')$ is not differentiable at $h' = 0$. To see this, first consider the case when the initial health stock $h$ is large relative to the health shock $\varepsilon$ so that $(1 - \delta_h) h + \varepsilon \geq 0$. This means conditional on drawing $\varepsilon$ at age $j$, even if medical spending is infinitesimal, there is still a strictly positive chance of survival. The optimal decision rules for
consumption and medical care must then satisfy the first-order conditions:

\[
\beta \Phi [h_{j+1}(\theta, \varepsilon)] R [h_{j+1}(\theta, \varepsilon)] E^{j+1} \{U' [c_{j+1}(\theta', \varepsilon')]\} = U' [c_j(\theta, \varepsilon)], \tag{2.4}
\]

\[
\beta \Phi' [h_{j+1}(\theta, \varepsilon)] s' [m_j(\theta, \varepsilon)] E^{j+1} \left\{ V^{j+1}(\theta', \varepsilon') + \frac{\partial V^{j+1}(\theta', \varepsilon')}{\partial \theta'} \Phi [h_{j+1}(\theta, \varepsilon)] \right\} = pU' [c_j(\theta, \varepsilon)], \tag{2.5}
\]

where

\[
\frac{\partial V^j(\theta, \varepsilon)}{\partial h} = U' [c_j(\theta, \varepsilon)] R' [h_{j+1}(\theta, \varepsilon)] s_{-1} + (1 - \delta_h) \frac{pU' [c_j(\theta, \varepsilon)]}{v' [m_j(\theta, \varepsilon)]}. \tag{2.6}
\]

Condition (2.4) is the Euler equation that governs consumption growth. The optimality condition for medical care is given by (2.5). Strict concavity of \( \Phi(m) \) ensures that \( m_j(\theta, \varepsilon) > 0 \).

Suppose now \((1 - \delta_h) h + \varepsilon < 0\). The relationship between \( m \) and \( h' \) is depicted in Figure 2.1. In this case, any \( m \in [0, m] \) would yield a non-positive level of health after treatment and hence zero probability of survival. The optimal demand for medical services is thus either zero or above \( m \). If \( m_j(\theta, \varepsilon) = 0 \) then survival to the next period is impossible and the first-order conditions (2.4) and (2.5) are no longer applicable.\(^6\) In this case, the current period is terminal for the agent and it is thus optimal to consume all resources available and have zero savings.\(^7\) If \( m_j(\theta, \varepsilon) > m_\) then the optimal decision rules must again satisfy (2.4) and (2.5).

Intuitively, zero would be the optimal choice if the alternative is either infeasible

\(^6\)Note that this corner solution cannot be ruled out even if \( \lim_{m \to 0} \Phi'(m) = \infty \). This is because the health shock enters into the accumulation equation (1.1) in an additive fashion.

\(^7\)Under full information, the fact that the agent has no hope to survive is publicly known. So even though the agent has an huge incentive to borrow, no one will lend him any loans.
or too costly to achieve.\textsuperscript{8} As in section 1.3, all the decision rules can be expressed in terms of the shock history.

\subsection{Competitive Equilibrium}

This section describes a competitive equilibrium for a stationary environment in which the life-cycle patterns of individual consumption, medical spending and savings are time-invariant. It follows that all the aggregate variables and the population structure are time-invariant in this environment.

Let $N(z^j)$ denote the measure of all age-$j$ agents with shock history $z^j$. The population measure evolves according to

\begin{equation}
N(z^j) = \pi_j(\varepsilon) \tilde{\Phi}[h_j(z^{j-1})] N(z^{j-1}), \quad \text{for } j = 1, \ldots, J, \tag{2.7a}
\end{equation}

with $z^j = (z^{j-1}, \varepsilon), \tilde{\Phi}[h_j(z^{j-1})] \equiv \Phi[h_j(z^{j-1})] / (1 + \gamma)$, and

\begin{equation}
N(z^0) = \pi_0(\varepsilon), \tag{2.7b}
\end{equation}

for $z^0 = \varepsilon \in \mathcal{E}$. The size of the total population is given by

$$
\mathcal{N} = \sum_{j=0}^{J} \sum_{z^j} N(z^j).
$$

All the annuitized savings are invested in physical capital. The capital market

\textsuperscript{8}The above discussion illustrates that it is optimal for the agents to end their lives under certain conditions. The same feature can be found in Grossman (1972), but the idea is very different. Grossman adopted a deterministic framework in which agents have full discretion over their health stock in each period. Death occurs when health falls below a certain level. Thus, given all the future prices and state of technology, one can actually choose the end of life at age $\theta$ which is not the case in here.
clears when aggregate demand equals aggregate supply; i.e.,

\[
k = \sum_{j=0}^{J} N(z^j) s_j(z^j). \tag{2.8}
\]

In the annuity market, perfect competition implies zero profit for each firm. This means all the returns from physical capital investment are distributed as benefits. For an age-\((j+1)\) agent with past history \(z^i\), the amount of benefits received is \(R[h_{j+1}(z^j)] s_j(z^j), \) for \(j = 0, ..., J - 1\). The size of this group is given by \(\Phi[h_{j+1}(z^j)] N(z^j)\). The amount of total annuity benefits can be obtained by summing across all types and all ages,

\[
\sum_{j=0}^{J-1} \sum_{z^j} \Phi[h_{j+1}(z^j)] N(z^j) R[h_{j+1}(z^j)] s_j(z^j).
\]

Let \(r\) be the (gross) rate of return from physical capital. The annuity market clears when the total benefits equal total returns from physical capital; i.e., \(rk\). Using (2.8), the market-clearing condition for annuity can be expressed as

\[
\sum_{j=0}^{J} N(z^j) s_j(z^j) \left\{ \Phi[h_{j+1}(z^j)] R[h_{j+1}(z^j)] - r \right\} = 0. \tag{2.9}
\]

The annuity return function is assumed to satisfy the following condition

\[
\Phi[h_{j+1}(z^j)] R[h_{j+1}(z^j)] = r, \tag{2.10}
\]

for all \(z^i\) and for \(j = 0, ..., J - 1\). In light of (2.9), this assumption rules out any cross-subsidization between different groups. Mortality risks are shared among those with the same medical history (and hence the same level of health). Within any group, the annuity can be viewed as an intergenerational transfer system.
In each period, the total annuitized wealth of all age-$j$ agents with history $z^j$ is distributed as annuity benefits to those with history $z^j$ in the previous period.

In the consumption good sector, firms hire labor and rent physical capital from the factor markets to produce output according to (1.2). Both goods and factor markets are assumed to be perfectly competitive, hence

$$r = \alpha A k^{\alpha - 1} l^{1-\alpha} + 1 - \delta_k,$$

$$w = (1 - \alpha) A k^\alpha l^{-\alpha}.$$  \hspace{1cm} (2.11) \hspace{1cm} (2.12)

The consumption good market clears when

$$\sum_{j=0}^{J} \sum_{z^j} N(z^j) \left[ c_j(z^j) + pm_j(z^j) \right] + \delta_k k = A k^\alpha l^{1-\alpha}.$$  \hspace{1cm} (2.13)

The labor market clears when the following holds,

$$l = \sum_{j=0}^{J} \sum_{z^j} N(z^j) e_j.$$  \hspace{1cm} (2.14)

A competitive equilibrium for this economy is defined as follows:

**Definition** A competitive equilibrium for the complete market economy consists of allocations $\varphi = \{c_j(z^{j-1}, \varepsilon), m_j(z^{j-1}, \varepsilon), s_j(z^{j-1}, \varepsilon), x_j(z^{j-1}) | z^{j-1} \in \mathcal{Z}^{j-1}, \varepsilon \in \mathcal{E} \}_{j=0}^{J}$, capital and labor inputs $\{k, l\}$, factor prices $\{r, w\}$, contingent claims prices $\{q_j(\varepsilon | z^{j-1}) | (z^{j-1}, \varepsilon) \in \mathcal{Z}^j \}_{j=0}^{J}$, population measures $\{N(z^j) | z^j \in \mathcal{Z}^j \}_{j=0}^{J}$, and an annuity return function $R(h)$ such that

1. Given the prices and returns, $\varphi$ solves the agents’ problem.

2. Factor inputs and prices, $\{k, l, r, w\}$ satisfy (2.11) and (2.12).
3. The annuity return function satisfies (2.10) and the prices of contingent claims are given by

\[ q_j(\varepsilon|z^{j-1}) = \pi_j(\varepsilon), \quad \text{for all } \varepsilon \in \mathcal{E} \text{ and } z^{j-1}. \]  \hspace{1cm} (2.15)

4. The population measures evolve according to (2.7a) and (2.7b).

5. All markets clear; i.e., (2.8), (2.13) and (2.14) hold.

2.3 Theoretical Analysis

When contingent claims are priced according to (2.15), all agents would perfectly smooth out their consumption across states. It follows that consumption is history-independent so that agents of the same age would consume the same amount. The optimal consumption profile is then characterized by the deterministic Euler equation. These are formally verified in the following lemma.

Lemma 2.1  Agents of the same age would have the same consumption; i.e., \( c_j(z^j) = c_j \) for all \( z^j \) and for all \( j \). The consumption profile \( \{c_j\}_{j=0}^J \) is characterized by the Euler equation

\[ \beta r U''(c_{j+1}) = U'(c_j). \]  \hspace{1cm} (2.16)

Proof. Condition (2.3) can be restated as

\[ \frac{\pi_j(\varepsilon) U''[c_j(z^{j-1}, \varepsilon)]}{q_j(\varepsilon|z^{j-1})} = \frac{\pi_j(\bar{\varepsilon}) U'[c_j(z^{j-1}, \bar{\varepsilon})]}{q_j(\bar{\varepsilon}|z^{j-1})}, \]

for any \( \varepsilon, \bar{\varepsilon} \in \mathcal{E} \). Condition (2.15) then implies perfect consumption-smoothing
across states; i.e., \( c_j(z^{j-1}, \varepsilon) = c_j(z^{j-1}, \tilde{\varepsilon}) \) for all \( j \). It follows that at age 0, \( c_0(\varepsilon) = c_0 \) for all \( \varepsilon \in \mathcal{E} \). Suppose \( c_n(z^n) = c_n \) for \( j = 0, \ldots, n \), with \( n < J \). Although equation (2.4) may not hold in some states, it must hold in the “healthy” state at age \( n \); i.e., when \( \varepsilon^n = 0 \). By the perfect consumption smoothing result, one can replace \( E_{n+1}\{U'[c_{n+1}(z^{n+1})]\} \) with \( U'[c_{n+1}(z^n)] \). Combining with (2.10), condition (2.4) then becomes

\[
\beta r U'[c_{n+1}(z^n)] = U'(c_n).
\]

This establishes that consumption at age \( (n + 1) \) depends on age alone, i.e. \( c_{n+1}(z^n) = c_{n+1} \). This completes the proof.

The next theorem provides a sufficient condition under which the competitive equilibrium is Pareto optimal. Thus, the complete-market economy can be used as a benchmark for welfare analysis.

**Theorem 2.2** A competitive equilibrium, which converges to a steady state with long-run gross interest rate \( r^* > 1 + \gamma \), is Pareto optimal.

**Proof.** See Appendix.

### 2.4 Findings

In the quantitative analysis, the calibration procedure described in section 1.4 is imposed on the complete-market economy. Table 2.1 compares the results obtained from the complete and incomplete-market economies under the same parameterization. Similar to the economy with no insurance, the complete-market economy could yield similar (but slightly lower) levels of life expectancy at all ages with fewer medical spending. Having an ideal market structure would
lower the share of medical spending in GDP by 1.9 percentage points in 1950 and 2.8 percentage points in 2001. These findings suggest a source of inefficiency in the incomplete-market economy. Table 2.1 also shows how the allocation of medical services differ under the two market structures. When compared to the incomplete-market economy in 1950, an ideal market structure would allocate more medical services to the elderly population. The contrary is true when it is compared to the incomplete-market economy in 2001 when public insurance is available for the elderly.

Most importantly, a large increase in medical spending over time is observed in the complete-market economy in which all contractual and institutional changes pertaining to health insurance are removed. This suggests that the increase is driven by technological improvement in medical treatment and rising incomes which are common in both complete-market and incomplete-market economies.

2.5 Welfare Analysis

In the welfare analysis, a consumption equivalent variation measure is used to quantify the potential welfare loss due to market incompleteness. The motivation for this measure is described below. Consider two agents who are ex ante identical except for one thing: the first consumer is born in the complete-market economy and faces the equilibrium allocation \((c, m) = \{c_j, m_j(z^j) | z^j \in Z^j\}_{j=0}^J\), while the second one appears in the incomplete-market economy and chooses to consume \((\bar{c}, \bar{m}) = \{\bar{c}_j(z^j), \bar{m}_j(z^j) | z^j \in Z^j\}_{j=0}^J\). Their expected lifetime utility are denoted by \(W(c, m)\) and \(W(\bar{c}, \bar{m})\), respectively.\(^9\) An equivalent variation

\(^9\)To be more precise, \(W(c, m)\) is equivalent to \(E^0[V^0(\theta, \varepsilon)]\) defined in (P2.1), and \(W(\bar{c}, \bar{m})\) is equivalent to the same expression defined in either (P1.1) or (P1.2) depending on the time period in question.
measure, $\phi$, is defined as follow:

$$W(c, m) = W(\phi\tilde{c}, \tilde{m}).$$

This above condition states that consumers born in the incomplete-market economy would be as well-off as those born in the complete-market economy if their consumption is increased by $\phi$ in every period and in every possible state. With the specific form of period utility function, this measure can be expressed as

$$\phi = \left[ \frac{W(c, m)}{W(\tilde{c}, \tilde{m})} \right]^{\frac{1}{1-\sigma}}.$$

Recall that in the model, one period takes five years. Thus the annual rate of change of consumption is $\Delta = (\phi^{0.2} - 1)$. The value of $\Delta$ is $0.28\%$ in 1950 and $0.76\%$ in 2001.

### 2.6 Concluding Remarks

In this chapter, a complete-market version of the economy presented in Chapter 1 is considered. Consumers in this economy can perfectly insure against any uncertainty in their consumption by trading a combination of state-contingent claims and actuarially fair annuities. The model economy is then calibrated using the same procedure as described in section 1.4.

There are three major results in this chapter. In the theoretical analysis, it is shown that Pareto optimality can be achieved in equilibrium under this market structure. Hence, the complete-market economy is a natural benchmark for welfare analysis. In the quantitative analysis, it is found that a large increase in medical spending over time can be obtained even in the first-best scenario. This
reinforces the conclusion obtained in Chapter 1. Finally, in the welfare analysis, the extent of welfare loss due to market incompleteness is quantified.
2.7 Appendix

The objective of the appendix is to prove Theorem 2.2. This is divided into two parts and a number of intermediate steps. The first part involves specifying the planner’s problem and the conditions under which a solution to this problem is Pareto optimal. The second part shows that for any competitive equilibrium allocation, one can construct a planner’s problem so that the equilibrium allocation is a solution. It also shows that the conditions for Pareto optimality are satisfied.

Throughout the appendix, the consumption and medical care of an age-$j$ agent at time $t$ with shock history $z^j$ are denoted by $c_t(z^j)$ and $m_t(z^j)$, respectively. Moreover, this agent is said to be of cohort $(t-j)$.

The Planner’s Problem

Let $\alpha_t = \{c_{t+j}(z^j), m_{t+j}(z^j) \mid z^j \in Z^j\}_{j=0}^J$ be an arbitrary consumption profile for an agent of cohort $\tau$. Conditional on being alive at age $j$ with past history $z^{j-1}$, the expected lifetime utility that the agent derived from $\alpha_t$ is

$$W^j(\alpha_t; z^{j-1}) = \sum_{\epsilon \in E} \pi_j(\epsilon) \left\{ U [c_{t+j}(z^j)] + \beta \Phi \left[ h_{t+j+1}(z^j) \right] W^{j+1}(\alpha_t; z^j) \right\},$$

where $z^j = (z^{j-1}, \epsilon)$ and

$$h_{t+j+1}(z^j) = i [m_{t+j}(z^j)] + (1 - \delta_h) h_{t+j}(z^{j-1}) + \epsilon,$$

for $j = 0, \ldots, J - 1$. The expected utility at the terminal age is

$$W^J(\alpha_t; z^{J-1}) = \sum_{\epsilon \in E} \pi_J(\epsilon) U [c_{t+J}(z^J)].$$
Let $U(\alpha_\tau)$ be the expected lifetime utility derived from this profile. For any initial old agents of generation $\tau = -1, \ldots, -J$,

$$U(\alpha_\tau) \equiv W^{-\tau}(\alpha_\tau; z^{-1}).$$

For all other generations,

$$U(\alpha_\tau) \equiv W^0(\alpha_\tau; z^{-1}).$$

Denote by $N_t(z^j)$ the measure of age-$j$ agents with shock history $z^j$ at time $t$. At time $t = 0$, the measures of various types of agents are given by

$$N_0(z^j) = \pi_j(\varepsilon)(1 + \gamma)^{-j},$$

where $z^j = (z^{-1}, \varepsilon)$ for all $j$, and $z^{-1}$ is an empty sequence. At time $t > 0$, the measure is defined recursively as

$$N_t(z^0) = \pi_0(\varepsilon)(1 + \gamma)^t, \quad \text{for } z^0 = \varepsilon,$$

$$N_t(z^j) = \pi_j(\varepsilon) \Phi[h_t(z^{j-1})] N_{t-1}(z^{j-1}),$$

for $j \geq 1$, $z^j = (z^{j-1}, \varepsilon)$ and

$$h_{t+1}(z^j) = i[m_t(z^j)] + (1 - \delta_h)h_t(z^{j-1}) + \varepsilon.$$

A feasible allocation at any time $T \geq 0$, denoted by $\phi_T$, consists of a set of allocation rules for all agents ever exist from time $T$ onwards, $\{\alpha_{t-j}\}_{t \geq T}$, and a sequence of aggregate savings starting from time $T$, $\{a_t\}_{t \geq T}$, that satisfy the
following for all $t \geq T$,

$$
\sum_{j=0}^{J} \sum_{z^j} N_t(z^j) [c_t(z^j) + p_t m_t(z^j)] + a_t = Ak_t^\alpha l_t^{1-\alpha},
$$

$$
k_{t+1} = a_t + (1 - \delta_k) k_t,
$$

$$
l_t = \sum_{j=0}^{J} \sum_{z^j} N_t(z^j) c_j,
$$

and the population dynamics specified by (2.17)-(2.20). The set of all feasible allocations at time $T$ is denoted by $\mathcal{A}_T$.

The current state of the economy is summarized by $S_T = (k_T, h_T, N_T)$, where $k_T$ denotes the predetermined level of capital stock, $h_T = \{h_T(z^j) | z^j \in Z^j\}_{j=0}^{J}$ summarizes the health level of all types of agents, and $N_T = \{N_T(z^j) | z^j \in Z^j\}_{j=0}^{J}$ summarizes the distribution of health shock sequence, $z^j$, in the population. At time 0, all agents (including the initial old) have the same level of health, $\overline{h} > 0$. The state of the economy can be expressed as $S_0 = (k_0, \overline{h}, N_0)$.

Let $\theta_\tau$ be the welfare weight that the planner assigns to agents of generation $\tau \geq -J$. The planner’s problem at time 0 is to choose a feasible allocation so as to maximize the welfare of all generations, i.e.,

$$
\tilde{V}_0(S_0) = \max_{\phi_0 \in \mathcal{A}_0} \sum_{\tau=-J}^{\infty} \theta_\tau (1 + \gamma)^\tau U(\alpha_\tau).
$$

In recursive form, the social planner’s value at time $t \geq 0$ is

$$
\tilde{V}_t(S_t) = \max_{\phi_t \in \mathcal{A}_t} \left\{ \sum_{j=0}^{J} \sum_{z^j} \tilde{\beta}(t, j) \theta_{t-j} N_t(z^j) U\left[c_t(z^j)\right] + \tilde{V}_{t+1}(S_{t+1}) \right\},
$$
where

$$\tilde{\beta}(t,j) = \begin{cases} 
\beta^t & \text{for } t - j < 0 \\
\beta^j & \text{for } t - j \geq 0.
\end{cases}$$

The subjective discount factor is time-dependent because the initial old agents discount their utilities at each age back to time 0, while all other agents discount their utilities back to their date of birth.

The first-order condition with respect to $c_t(z^j)$ is

$$\theta_{t-j}\tilde{\beta}(t,j)U''[c_t(z^j)] = \frac{\partial \tilde{V}_{t+1}(S_{t+1})}{\partial k_{t+1}}, \quad (2.21)$$

for all $z^j$. Since $(2.21)$ is independent of $z^j$, the social planner would allocate consumption so that all living agents from the same cohort would have the same consumption; i.e., $c_t(z^j) = c_{t,j}$ for all $z^j$ and for all $t$. For a particular cohort and ability level, the consumption profile is characterized by the typical Euler equation

$$\beta r_{t+1}U'(c_{t+1,j+1}) = U'(c_{t,j}). \quad (2.22)$$

For the initial old agents of generation $-j$ who are alive at time $t$, $(2.21)$ implies

$$\theta_{-j}U'(c_{t,t+j}) = \theta_0 U'(c_{t,t}), \quad (2.23)$$

for $j = 1, ..., J$. For agents of all other generations, $(2.21)$ implies

$$\theta_{t-j}\beta^j U'(c_{t,j}) = \theta_t U'(c_{t,0}), \quad (2.24)$$

for all $t \geq 0$ and for all $j$.

Next, consider the allocations of medical care. If $(1 - \delta_h) h_t(z^{i-1}) + \varepsilon > 0$ and
\(z^j = (z^{j-1}, \varepsilon)\), then allocating any nonnegative amount of medical care would yield \(N_{t+1}(z^j) > 0\). The optimal allocation \(m_t(z^j)\) must satisfy the first-order condition:

\[
\frac{\partial \tilde{V}_{t+1}(S_{t+1})}{\partial h_{t+1}(z^j)} \frac{\partial h_{t+1}(z^j)}{\partial m_t(z^j)} + \sum_{\varepsilon \in E} \frac{\partial \tilde{V}_{t+1}(S_{t+1})}{\partial N_{t+1}(z^j, \varepsilon)} \frac{\partial N_{t+1}(z^j, \varepsilon)}{\partial m_t(z^j)} = \frac{\partial \tilde{V}_{t+1}(S_{t+1})}{\partial k_{t+1}} p_{N_t}(z^j),
\]

(2.25)

where

\[
\frac{\partial h_{t+1}(z^j)}{\partial m_t(z^j)} = i'[m_t(z^j)],
\]

\[
\frac{\partial N_{t+1}(z^j, \varepsilon)}{\partial m_t(z^j)} = \pi_j(\varepsilon) N_t(z^j) \Phi'[h_{t+1}(z^j)] i'[m_t(z^j)].
\]

By allocating an additional unit of medical services to an agent with shock sequence \(z^j\), social welfare would increase due to (i) an improvement in the end-of-period health status of this agent, (ii) an increase in the number of survivors in the next period. With the help of the Envelope Theorem, and after some tedious algebras, (2.25) can be expressed as

\[
\beta \sum_{\varepsilon \in E} \pi_{j+1}(\varepsilon) \left\{ V_{t+1}(z^j, \varepsilon) + Q_{t+1}(z^j, \varepsilon) + (1 - \delta) p_{t+1} \frac{\Phi[h_{t+1}(z^j)]}{\Phi'[h_{t+1}(z^j)]} \frac{U'(c_{t+1,j+1})}{i'[m_{t+1}(z^j)]} \right\} = \frac{p_t U'(c_{t,j})}{\Phi'[h_{t+1}(z^j)]} \Phi'[h_{t+1}(z^j)] i'[m_t(z^j)].
\]

(2.26)

The functions \(V_t(z^j)\) and \(Q_t(z^j)\) are defined recursively as

\[
V_t(z^j) = U(c_{t,j}) + \beta \Phi[h_{t+1}(z^j)] \sum_{\varepsilon \in E} \pi_{j+1}(\varepsilon) V_{t+1}(z^j, \varepsilon),
\]
with \( V_t(z^j) = U(c_{t,j}) \) for all \( t \), and

\[
Q_t(z^j) = U'(c_{t,j}) B_t(z^j) + \beta \Phi \left[ h_{t+1}(z^j) \right] \sum_{\varepsilon \in \mathcal{E}} \pi_{j+1}(\varepsilon) Q_{t+1}(z^j, \varepsilon), 
\]

\[ (2.27) \]

and \( Q_t(z^j) = -U'(c_{t,J}) c_{t,J} \) for all \( t \). The variable \( B_t(z^j) \) is defined by

\[
B_t(z^j) \equiv w_t e^j - c_{t,j} - p_t m_t(z^j),
\]

If \( (1 - \delta_h) h_t(z^{j-1}) + \varepsilon < 0 \), then define \( m_t(z^j) \) such that

\[
i \left[ m_t(z^j) \right] + (1 - \delta_h) h_t(z^{j-1}) + \varepsilon = 0.
\]

By the strict concavity of \( i(\cdot) \), there exists a unique \( m_t(z^j) > 0 \) that solves the above equation. It is never optimal for the social planner to allocate any medical services between \( 0 \) and \( m_t(z^j) \) to this group. The reason is any \( m \) within this range would not extend the agents’ lives and thus add nothing to the social welfare. So the optimal allocation is either above \( m_t(z^j) \) or zero.

The rest of the section proceeds under the assumption that the above economy eventually converges to a stationary environment in which all individual and per-capita variables are time-variant, while all aggregate variables are growing at the population growth rate \( (\gamma) \). The following lemma states the condition under which \( \tilde{V}_0(S_0) \) is bounded.

**Lemma 2.3** If \( \sum_{\tau=-J}^{\infty} \theta_\tau (1 + \gamma)^\tau < \infty \), then the value function of the social planner at time 0 is bounded, i.e. \( \tilde{V}_0(S_0) < \infty \) for all \( S_0 \).

**Proof.** Let \( \{c_{t,j}\}_{t\geq 0, j=0,\ldots,J} \) be a consumption sequence that solves the social planner’s problem. Under the convergence assumption, \( c_{t,j} \to c_j \) as \( t \to \infty \). By feasibility,
individuals’ consumption must be finite; i.e., $c_j < \infty$ for all $j$. This means there exists a positive number $\mathcal{U} < \infty$ such that $U(c_j) < \mathcal{U}$ for all $j$. In particular, one can pick $\mathcal{U}$ such that $U(c_{t,j}) < \mathcal{U}$ for all $t$. In the presence of discounting and survival probability, the expected lifetime utility of any cohort must be strictly less than $JU$. Thus, the social planner’s value at time 0 must satisfy

$$
\tilde{V}_0(S_0) \equiv \sum_{\tau=-J}^{\infty} \theta_{\tau} (1 + \gamma)^{\tau} U(\alpha_{\tau}) < \sum_{\tau=-J}^{-1} \theta_{\tau} (1 + \gamma)^{\tau} U(\alpha_{\tau}) + JU \sum_{\tau=0}^{\infty} \theta_{\tau} (1 + \gamma)^{\tau}.
$$

The infinite sum is finite if $\sum_{\tau=-J}^{\infty} \theta_{\tau} (1 + \gamma)^{\tau}$. This completes the proof.

\textit{Pareto Optimality}

The objective of this section is to establish a connection between the solutions of the social planner’s problem and Pareto optimal allocations.

A feasible allocation at time 0, $\phi_0 \in A_0$, is optimal if there does not exist other feasible allocation $\tilde{\phi}_0$ which increases the welfare of some agents without reducing the welfare of others, i.e.

$$U(\tilde{\alpha}_\tau) \geq U(\alpha_\tau),$$

for $\tau \geq -J$, with strict inequality hold for some cohort $\tau$.

\textbf{Lemma 2.4} \textit{If} $\sum_{\tau=-J}^{\infty} \theta_{\tau} (1 + \gamma)^{\tau} < \infty$, \textit{then any solution of the planner’s problem is Pareto optimal.}

\textbf{Proof.} Let $\phi_0$ be a solution of the planner’s problem and suppose the contrary that it is not Pareto optimal. Then there exists an alternative allocation $\tilde{\phi}_0 \in A_0$ such that (2.28) are true with strict inequality hold for some $\tau$. Multiplying both
sides of (2.28) by \( \theta_t (1 + \gamma)^T > 0 \), and summing across all cohorts gives

\[
\sum_{\tau = -J}^{\infty} \theta_t (1 + \gamma)^T U(\tilde{\alpha}_\tau) > \tilde{V}_0(S_0) .
\] (2.29)

The comparison is valid if \( \sum_{\tau = -J}^{\infty} \theta_t (1 + \gamma)^T < \infty \). First, feasibility requires that \( U(\tilde{\alpha}) < \infty \). Second, \( \tilde{V}_0(S_0) < \infty \) by Lemma 2.3. But then (2.29) implies that \( \phi_0 \) is not a solution of the planner’s problem, hence a contradiction.

\textbf{Proof of Theorem 2.2}

The proof is made up of two parts: (i) It is shown that given any competitive equilibrium allocation, one can construct a planner’s problem so that the equilibrium allocation is a solution. This involves constructing a set of weights, \( \hat{\theta} = \{\hat{\theta}_t|\tau \geq -J\} \), so that the equilibrium allocation solves the planner’s problem under these weights. This part is further divided into a number of intermediate steps. (ii) It is verified that \( \hat{\theta} \) satisfy the condition in Lemma 2.4 and hence the equilibrium is Pareto optimal.

\textbf{Step 1}  Let \( \{\tilde{c}_{t,j}, \tilde{m}_t(z^j), \tilde{s}_t(z^j), \tilde{x}_t(z^j) | z^j \in Z^j\}_{t \geq 0}^{j=0,...,J} \) be an equilibrium allocation.

First, construct the sequence of equilibrium population measures, \( \{\tilde{N}_t(z^j) | z^j \in Z^j\}_{t \geq 0}^{j=0,...,J} \) according to (2.7a)-(2.7b). Second, construct the sequence of equilibrium inputs and interest rate, \( \{\tilde{k}_{t+1}, \tilde{I}_t, \tilde{r}_t\}_{t=0}^{\infty} \), according to (2.8), (2.11) and (2.14), respectively. Given the equilibrium interest rate \( \tilde{r}_{t+1} \), conditions (2.16) and (2.22) are identical.
Step 2 Next construct a set of weights as follow: (i) For the initial old agents of generation $-j$, 

$$\hat{\theta}_{-j} = \frac{U''(\tilde{c}_{0,0})}{U''(\tilde{c}_{0,n})} \hat{\theta}_0 > 0,$$  

(2.30)

for $j = 1, ..., J$. (ii) For all other generations,

$$\hat{\theta}_t = \prod_{i=0}^{j-1} \left( \frac{1}{n_{t-i}} \right) \frac{U''(\tilde{c}_{t-j,0})}{U''(\tilde{c}_{t,0})} \hat{\theta}_{t-j} > 0,$$  

(2.31)

for all $j$ and $t - j \geq 0$, with $\hat{\theta}_0 = 1$.

Step 3 Expression (2.30) implies that for any agent of generation $-j$, $j > 0$, at time 0,

$$\hat{\theta}_{-j} U''(\tilde{c}_{0,j}) = \hat{\theta}_0 U''(\tilde{c}_{0,0}).$$

Updating this to time $t$ with the help of (2.22) gives

$$\Rightarrow \hat{\theta}_{-j} U''(\tilde{c}_{t,t+j}) = \hat{\theta}_0 U''(\tilde{c}_{t,t}),$$

which coincides with (2.23). Hence, consumption for the initial old agents in a decentralized economy would coincide with that determined by a social planner.

Step 4 Given (2.22) and (2.31), one can derive the following

$$\hat{\theta}_{t-j} \beta^j U''(\tilde{c}_{t,j}) = \hat{\theta}_t U''(\tilde{c}_{t,0}),$$

which coincides with (2.24) for $t - j \geq 0$. 

Step 5 Next, we have to show that (2.5) is identical to (2.26). Using (2.5), (2.6) can be rewritten as

\[
\frac{\partial V_t(z^j)}{\partial h_t(z^{j-1})} = U'(c_{t,j}) R' [h_t(z^{j-1})] s_{t-1}(z^{j-1}) + (1 - \delta_h) \frac{p_t U'(c_{t,j})}{i'[m_t(z^j)]}
\]

Substituting this back into (2.5) and differentiating (2.10) with respect to \( h' \) gives

\[
\beta \sum_{\varepsilon \in \mathcal{E}} \pi_{j+1}(\varepsilon) \left\{ V_{t+1}(z^j, \varepsilon) - U'(c_{t+1,j+1}) R [h_{t+1}(z^j)] s_t(z^j) 
\right.
+ (1 - \delta_h) p_{t+1} \frac{\Phi [h_{t+1}(z^j)]}{\Phi' [h_{t+1}(z^j)]} \frac{U'(c_{t+1,j+1})}{i'[m_{t+1}(z^j, \varepsilon)]} \right\}
\]

\[
= \frac{p_t U'(c_{t,j})}{\Phi' [h_{t+1}(z^j)]} \frac{\Phi [h_{t+1}(z^j)]}{i'[m_t(z^j)]} \tag{2.32}
\]

for \( j = 0, ..., J - 1 \). Comparing between (2.32) and (2.26), one can see that it suffice to show that

\[
\sum_{\varepsilon \in \mathcal{E}} \pi_{j+1}(\varepsilon) \tilde{Q}_{t+1}(z^j, \varepsilon) = -U'(\tilde{c}_{t+1,j+1}) R [\tilde{h}_{t+1}(z^j)] \tilde{s}_t(z^j), \tag{2.33}
\]

where \( \tilde{s}_t(z^j) \) is defined according to (2.27) based on the equilibrium allocation.

An induction argument is used to show that (2.33) holds. Since this argument holds for all time \( t \), the time subscript is removed to ease the burden of notations. At age \( J - 1 \), the LHS of (2.33) is just \(-U'(\tilde{c}_J) \tilde{c}_J\). The RHS would gives the same expression by applying the following

\[
R [\tilde{h}_J(z^{J-1})] \tilde{s}(z^{J-1}) = \tilde{c}_J.
\]
Next, suppose (2.33) is true for \( j = n + 1, \ldots, J - 1 \). Then (2.27) becomes

\[
\tilde{Q}(z^n) = U'(\tilde{c}_n) \tilde{B}(z^n) - \beta \Phi \left[ \tilde{h}'(z^n) \right] R \left[ \tilde{h}'(z^n) \right] U'(\tilde{c}_{n+1}) \tilde{s}(z^n)
\]

\[
= \quad U'(\tilde{c}_n) \left[ \tilde{B}(z^n) - \tilde{s}(z^n) \right]
\]

\[
= \quad -U'(\tilde{c}_n) \left\{ R \left[ \tilde{h}'(z^{n-1}) \right] \tilde{s}(z^{n-1}) + x(z^n) \right\}.
\]

The second line follows by using (2.10) and then (2.16). The third line follows from the individual’s budget constraint at age \( n \). Summing the above expression across all possible states and using the budget constraint for contingent claims yield

\[
\sum_{\varepsilon \in \varepsilon} \pi_{n+1}(\varepsilon) \tilde{Q}(z^n, \varepsilon) = -U'(\tilde{c}_{n+1}) R \left[ \tilde{h}'(z^n) \right] \tilde{s}(z^{n-1}).
\]

Hence, (2.33) is true for \( j = n \) and it follows that the equilibrium allocation of medical care would satisfy (2.26).

**Step 6** The economy-wide resources constraint is satisfied in every period by (2.13). This completes the first part of the proof.

Since the competitive equilibrium converges to a steady state in the long-run, this means for \( j = 0, \ldots, J, \tilde{c}_{t,j} \to \tilde{c}_j \) as \( t \to \infty \). Hence, for \( T > 0 \) sufficiently large, (2.31) becomes

\[
\hat{\theta}_{T+\tau} = \left( \frac{1}{r^*} \right)^\tau \hat{\theta}_T,
\]

for \( \tau > 0 \), where \( r^* \) denotes the long-run interest rate. Hence,

\[
\sum_{\tau=-J}^{\infty} \hat{\theta}_\tau (1 + \gamma)^\tau = \sum_{\tau=-J}^{T-1} \hat{\theta}_\tau (1 + \gamma)^\tau + \hat{\theta}_T \sum_{\tau=0}^{\infty} \left( \frac{1 + \gamma}{r^*} \right)^\tau.
\]

The infinite sum is defined as \( r^* > 1 + \gamma \).
2.8 Figures and Tables

Figure 2.1: Relationship between End-of-Period Health and Medical Care.

\[ h' = i(m) + (1 - \delta_h)h + \epsilon \]
Table 2.1
Findings: Complete vs Incomplete Markets.

<table>
<thead>
<tr>
<th></th>
<th>1950 Complete</th>
<th>1950 Incomplete</th>
<th>2001 Complete</th>
<th>2001 Incomplete</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Medical Spending</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% in GDP</td>
<td>1.8%</td>
<td>3.7%</td>
<td>9.6%</td>
<td>12.4%</td>
</tr>
<tr>
<td>% in PCE</td>
<td>2.6%</td>
<td>4.8%</td>
<td>13.7%</td>
<td>16.6%</td>
</tr>
<tr>
<td><strong>Share of medical spending by Age</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age 45+</td>
<td>56.2%</td>
<td>55.6%</td>
<td>57.7%</td>
<td>61.8%</td>
</tr>
<tr>
<td>Age 65+</td>
<td>15.6%</td>
<td>12.0%</td>
<td>17.9%</td>
<td>23.9%</td>
</tr>
<tr>
<td>Age 75+</td>
<td>5.0%</td>
<td>2.9%</td>
<td>6.3%</td>
<td>10.6%</td>
</tr>
<tr>
<td><strong>Share of Population by Age</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age 45+</td>
<td>51.8%</td>
<td>51.9%</td>
<td>54.4%</td>
<td>54.5%</td>
</tr>
<tr>
<td>Age 65+</td>
<td>17.8%</td>
<td>17.8%</td>
<td>21.3%</td>
<td>21.4%</td>
</tr>
<tr>
<td>Age 75+</td>
<td>7.6%</td>
<td>7.6%</td>
<td>10.4%</td>
<td>10.5%</td>
</tr>
<tr>
<td><strong>Life Expectancy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age 25</td>
<td>47.5</td>
<td>47.7</td>
<td>51.6</td>
<td>51.8</td>
</tr>
<tr>
<td>Age 45</td>
<td>28.8</td>
<td>28.9</td>
<td>32.8</td>
<td>32.9</td>
</tr>
<tr>
<td>Age 65</td>
<td>14.6</td>
<td>14.7</td>
<td>17.6</td>
<td>17.8</td>
</tr>
</tbody>
</table>
Chapter 3
Suburbanization and the Automobile

3.1 Introduction

Suburbanization has been observed in cities throughout the U.S. since the 1870’s when street railways were first implemented. But technological progress in transportation made its biggest contributions during the twentieth century with the invention of the automobile and later the modern highway system. The adoption of the car as the dominant form of transportation in the U.S., combined with rising real income levels, encouraged movement to less dense areas where housing was more affordable. The goal here is to assess, quantitatively, the relationship between the invention and diffusion of the automobile and suburbanization.

3.1.1 Suburbanization

Measuring Suburbanization Suburbanization is defined as the increased dispersion of urban population over land area.\(^1\) To measure the extent of suburban-

\(^1\)Hereafter, the terms “suburbanization”, “urban decentralization”, and “urban sprawl” refer to the same phenomenon.
ization, it is standard to assume that population density changes at a constant rate as one moves further from the city center. Thus researchers have adopted the following functional form to relate population density, \( d \), to distance from the city center, \( x \),

\[
d = ae^{-bx}.
\]  

(3.1)

The parameter, \( a \), is an estimate of the density (people/square mile) at the city center. The density gradient, \( b \), measures the rate of change of the density as the distance from the city center increases.\(^2\) The results from estimating equation (3.1) for Chicago and Atlanta are provided in Tables 3.1 and 3.2, respectively. Observe that Chicago is more densely populated than Atlanta, this can be seen by comparing the values of \( a \) across the two cities. The density gradient for Atlanta was much steeper than for Chicago in 1900. This can also be seen in Figures 3.1 and 3.2 which show the population density functions for Atlanta in 1900, 1950, and 1970 and Chicago in 1900 and 1950. The amount of suburbanization that occurred over a period of time is measured by the percentage decline in the population density gradient. In Chicago, the density gradient fell by 48\% over the period 1900-50. In Atlanta, it decreased by 39\% during the same period and 77\% over the period 1900-70.

\textbf{Facts} Suburbanization has been a widely observed phenomenon both in the U.S. and in countries across the world. Decentralization of American cities was first observed in the second half of the nineteenth century.\(^3\) During the period 1910-1970, declining population density gradients have been observed for almost all U.S. cities. It is during this period that suburbanization in the U.S. occurred

\(^2\)Exponential population density functions were first introduced in Clark (1951). For more recent studies, see McDonald (1989) and Mills and Tan (1980).

\(^3\)See Mieszkowski and Mills (1993) for a survey of suburbanization.
most intensively. This includes cities with either rising or declining populations over the period. Estimates of the population density gradient \( b \) are shown in Table 3.3 for forty-one U.S. cities that were defined as metropolitan districts in 1900. The table shows the density gradients and the percentage change in the density gradient over the period 1900-1970. The density gradients have been declining for the sample as a whole from 1910 to 1970 with the largest decline occurring in the 1940’s. The density gradients declined by 77% over the period 1900-1970.

**Theories of Suburbanization** A variety of explanations of decentralization have been proposed. One popular theory is that the rich move to suburbs in order to avoid the disamenities of the inner cities, such as high crime rates and poor schools.\(^4\) While these factors may be useful for explaining suburbanization observed in particular cities during particular time periods, they cannot explain the more general trend of decentralization observed in the U.S. since 1870 and in other countries.

Another theory of suburbanization focuses on the impact of technological progress in transportation which has reduced the costs of commuting over time. As the cost of commuting falls, higher income groups move further from the city center to enjoy more modern housing, more space, and/or more attractive communities. Beginning in the 1830’s, several innovations in mass transportation, including omnibus, commuter railroads and streetcars were introduced into American cities. These improvements enhanced the mobility of city dwellers and helped to expand large cities like New York, Boston and Philadelphia.\(^5\) According to historians such as Warner (1962) and Ward (1971), the first major movement of

\(^4\)See for instance, Cullen and Levitt (1999), Mills and Lubuele (1999) and the references therein.

\(^5\)See Taylor (1966) for a detailed historical account on these innovations.
wealthy individuals to the suburbs occurred during the 1850’s and 1860’s and was caused by the introduction of the streetcar. Yet, by far the biggest breakthrough in transportation came with the invention of the automobile. The extensive suburbanization observed during the twentieth century occurred simultaneously with the adoption of the automobile by Americans. Included in this adoption was the adaption of American cities to private vehicle commuting. This theory of suburbanization as occurring concurrently with technological progress in transportation, is developed and analyzed in the work of LeRoy and Sonstelie (1983). A more recent discussion of the theory can be found in Glaeser and Kahn (2004).

3.1.2 Automobiles

The structure of automobile ownership in the U.S. changed immensely throughout the twentieth century. In 1906 approximately one tenth of Americans owned a car. In fact, automobiles were thought of more as a toy for the rich than as a realistic mode of transportation. By the mid 1930’s, however, more than 44% of U.S. families owned a car. By 1995, 92% of American families owned at least one car and 59% owned two or more. The increase in car use and ownership can be seen in Figure 3.3, which shows the number of registered automobiles per capita in the U.S. from 1900 to 1993 and in Table 3.4 which presents the number of registered vehicles per person age twenty to sixty-four during 1900-1970. Figure 3.4 shows how car-ownership evolved over the second half of the century.

What caused the rise in car-ownership? Both ownership and number of cars

---


owned are increasing with household income. As shown in Figure 3.5, car-ownership ranges from less than 50% for the poorest income group to over 90% for the richest, over the period 1952-1965. Overall ownership increased from 65% to 74% during this period, largely due to the increase in the second and third quintiles.

There is a negative correlation between car-ownership and prices. The time cost of purchasing a new car fell by approximately 98% during the period 1906-2000. In 1906 a worker earning the average wage would have to work approximately 453 weeks or more than 8 years to acquire enough earnings to afford to buy a new car. By 1920 the average wage earner could purchase a new car with approximately 1.6 years of earnings. In 2000 the average worker could afford a new car with only 16 weeks of earnings. The time cost, in logarithms, can be seen in Figure 3.6. The figure depicts the time cost for an individual earning the average wage.

As prices fall not only are families at lower and lower income levels able to afford automobiles but the purchasing of two and eventually more than two vehicles becomes a reasonable expenditure. Owning two cars then gives a two-headed household more freedom to move further away from central locations. This argument is supported by Figure 3.7 which depicts the positive relationship between car-ownership and distance from the city center. The same figure also shows

---

8 Source: *Survey of Consumer Finances*, University of Michigan.

9 The time cost of purchasing a car is obtained by dividing the price of automobiles by labor income. The value for 1906 is then normalized to one. Data on automobile prices are collected from various sources: For the period 1906-1940, data reported by Raff and Trajtenberg (1995) are used; for 1947-1983, data are obtained from Gordon (1990); and for 1967-2000, the data are taken from Ward (2002). The Raff and Trajtenberg (1995) and Gordon (1990) price series are of quality-adjusted prices. The series from Ward (2002) is the price for a comparable car. The quality-adjusted price of a new car decreased by 85% since 1906. The price decreased at an average annual rate of approximately 2% throughout the period 1906-2000. The fastest rate of decrease occurred during the period 1906-40 with an average rate of 5.5%. Labor income is the average hourly wages of workers. This is computed with the help of the US real wage indices reported in Williamson (1995). Here it is assumed that agents work 40 hours a week.

10 Source for Figures 3.7 and 3.8: *Survey of Consumer Finances*, University of Michigan.
that families living further away from the city center are more likely to own more than one car, than families living closer to the center. Another piece of evidence is Figure 3.8, which shows car-ownership by city versus suburban location. When compared to those living in the central cities, residents in suburban areas have a higher rate of ownership and own more cars. This phenomenon becomes more predominant over time. By 1961, only 6% of those living in the central cities of the 12 largest SMSA’s had two or more cars, as compared to 17% in the suburbs. Ten years later, the two figures rose to 15% and 41%, respectively.

3.1.3 The Analysis

In order to assess the relationship between rising car-ownership and suburbanization, a general equilibrium model is developed. In the model an agent can choose his mode of transportation for commuting to work and his residential site. An agent can either take a bus, which is publicly owned and operated, or purchase and use a car. The bus in the model serves as a proxy for all forms of public transportation that are relatively cheaper and slower than the automobile. When agents make their location choices, they take into consideration two factors: the cost of commuting and the cost of housing. The desire to save on commuting costs pulls them closer to the employment center. This in turn generates a large demand for housing around the center and bids up rents. Optimal location choice thus involves a balance between the two. By owning a car, one can spend less time commuting from a given location. This induces agents to spread to neighboring suburbs and enjoy larger living spaces. But not everyone will choose to own a car. The price of a car serves as a fixed cost that screens out those with low incomes. As income rises, car prices decline, and the cost of commuting by car relative to taking the bus rises, automobiles become more affordable and attractive, and this
promotes suburbanization.\textsuperscript{11}

The model is first analyzed theoretically. It is shown that, in equilibrium, a strictly positive relationship between income and location exists. Thus there is a one-to-one mapping between income levels and locations. This feature of the model facilitates in solving for an equilibrium rent function which clears the land market at each location. In addition, there exists a unique threshold income level. Agents with income below this level choose to be bus-users while those above it become car-owners.

The theoretical analysis is followed by a quantitative portion in which the model is calibrated to match data on the changes in car-ownership and suburbanization over time. The results show that the model is able to account for 86\% of car-ownership in 1970 and 77\% of suburbanization which occurred between 1910 and 1970. The model’s predictions for commuting costs are also shown to be inline with commuting costs implied by the data.

This paper is most closely related to LeRoy and Sonstelie (1983). There are two major differences between the two works though. First, LeRoy and Sonstelie considered a framework in which agents are homogeneous. Thus they do not attempt to capture the relationship between income and location or income and car-ownership. Second, LeRoy and Sonstelie did not attempt to match their model to the data.

\textsuperscript{11}This paper abstracts from the location choices made by firms, and thus the decentralization of employment. See, for instance, Lucas and Rossi-Hansberg (2002) for a theoretical analysis of a model where identical firms and households choose their locations simultaneously.
3.2 The Model

3.2.1 The Environment

Consider a linear city located on the positive half of the real line. The city is of unit width so that the density of land at each location is equal to one. The upper boundary of the city is determined endogenously. Land beyond this boundary is used for agriculture. All production activities take place at the origin, or the city center and firms do not use land for production. The city is inhabited by a continuum of agents. The size of the total population is normalized to unity. Each agent is characterized by an ability, $\lambda$, drawn from a distribution $F(\lambda)$ defined on a bounded support $[\lambda_{\min}, \lambda_{\max}]$. The mean of the ability distribution is normalized to unity.

There are three types of goods in the economy: final goods, automobiles, and land. Land in the city is owned by a landowner who collects all the rent and spends it on consumption. During each period, the agents must commute from their residential location to their workplace at the origin. There are two modes of transportation in the city: car and bus.

3.2.2 The Bus-user’s Problem

If an agent with ability $\lambda$ chooses to commute by bus, then he chooses consumption in goods, $c$, consumption in land services, $l$, and location, $x$, to solve the static problem (P3.1), taking the rent function $q(\cdot)$ as given.

$$V^b(\lambda) = \max_{c,l,x} \{\alpha \ln c + (1 - \alpha) \ln l\}$$  \hspace{1cm} (P3.1)
subject to

\[ c + q(x)l = w\lambda - t(w\lambda, x), \]

\[ x \geq 0, \]

where \( q(x) \) is the rent at location \( x \), \( w \) is the market wage for an efficiency unit of labor, and \( t(w\lambda, x) \) is the cost for an agent with ability \( \lambda \) to travel distance \( x \) by bus. The last inequality restricts the location choice of bus-users to the positive half of the real line. The transportation cost function, \( t(\cdot) \), is assumed to be (i) twice continuously differentiable, (ii) increasing in both arguments, and (iii) satisfying \( t(w\lambda, 0) > 0 \) for all \( w\lambda \geq 0 \). The second assumption implies that the time cost of commuting is increasing in income. Since agents spend all their non-commuting time at work, those with higher wages have a higher opportunity cost of commuting. The third assumption implies that the cost function includes a fixed cost that is independent of distance.

Conditional on any given location \( x \geq 0 \), expenditures on goods and land by a bus-user with ability \( \lambda \) are given by

\[ c_b(\lambda, x) = \alpha [w\lambda - t(w\lambda, x)] \quad (3.2) \]

and

\[ q(x)l_b(\lambda, x) = (1 - \alpha) [w\lambda - t(w\lambda, x)]. \quad (3.3) \]

Let \( x_b(\lambda) \) denote the agent’s optimal location choice. This location choice is
characterized by the first-order condition\footnote{The nonnegativity constraint on location is binding for a measure 0 of agents since each location $x \geq 0$ has a measure 0 of land. Hence only interior solutions exist.}

\[-t_2(w\lambda, x) = q'(x) l_b(\lambda, x).\] (3.4)

By moving closer to the origin, the agent can reduce his transportation costs but this gain is balanced by an increase in rent. This implies that the equilibrium rent function is decreasing. Combining (3.3) and (3.4) gives

\[
\frac{q'(x)}{q(x)} = \frac{-t_2(w\lambda, x)}{(1 - \alpha) [w\lambda - t(w\lambda, x)]}.\] (3.5)

By the Implicit Function Theorem, if $q(\cdot)$ is twice continuously differentiable, then $x_b(\lambda)$ exists and is continuously differentiable. The following lemma states the condition under which $x_b(\lambda)$ is strictly increasing.

**Lemma 3.1** Given any twice continuously differentiable function $q(\cdot)$, the optimal location choice function $x_b(\lambda)$ is strictly increasing in $\lambda$ if and only if the income elasticity of housing demand exceeds the income elasticity of marginal commuting cost, or

\[
\frac{w\lambda t_{12}(w\lambda, x)}{t_2(w\lambda, x)} < \frac{w\lambda [1 - t_1(w\lambda, x)]}{w\lambda - t(w\lambda, x)}.\] (3.6)

**Proof.** See Appendix. ■

Since the rich (those with high ability) have a higher land demand than the poor (those with low ability) at any location, the former benefit by living further away from the city center where the rent is lower. However, the rich also have a higher time cost than the poor. This induces them to live closer to the city.
center. The location choice function is monotonically increasing if and only if the first effect dominates. If \( x_b(\lambda) \) is strictly monotonic, then any location \( x = x_b(\lambda) \) will be inhabited by bus-users with ability \( \lambda \) alone. Positive consumption requires that

\[
w\lambda \geq t[ w\lambda, x_b(\lambda) ] .
\]

### 3.2.3 The Car-owner’s Problem

If an agent with ability \( \lambda \) chooses to own a car, then he solves the static problem (P3.2), taking as given the rent function \( q(\cdot) \).

\[
V^c(\lambda) = \max_{c,l,x} \{ \alpha \ln c + (1 - \alpha) \ln l \} \quad \text{(P3.2)}
\]

subject to

\[
c + q(x) l = w\lambda - \tau(w\lambda, x) - p_c,
\]

\[
x \geq 0,
\]

where \( p_c \) denotes the price of a car, \( \tau(w\lambda, x) \) is the cost of traveling distance \( x \) by car. The transportation cost function \( \tau(w\lambda, x) \) is assumed to share the same properties as \( t(w\lambda, x) \). When comparing to the bus, a car takes less time to travel the same distance, i.e. \( \tau_2 < t_2 \), for all \( (w\lambda, x) \), but costs more to use, so that

\[
\tau(w\lambda, 0) + p_c > t(w\lambda, 0) \quad \text{(3.7)}
\]

holds for all \( w\lambda \).

Given (3.7), bus-users will always reside closer to the origin than car-owners. To see this consider an agent with ability \( \lambda \) who lives at the origin. If he owns a
car, his transportation expenses are $\tau(w\lambda, 0) + p_c$, which are higher than $t(w\lambda, 0)$, the cost of taking a bus. This means the agent would have a higher net income (net of transportation expenses) by switching to take a bus. Since it is never optimal to locate at the origin and to own a car, car-owners will reside at locations $x > 0$.

Conditional on any given location $x > 0$, expenditures on goods and land are given by

$$c_c(\lambda, x) = \alpha [w\lambda - \tau(w\lambda, x) - p_c], \quad (3.8)$$

$$q(x) l_c(\lambda, x) = (1 - \alpha) [w\lambda - \tau(w\lambda, x) - p_c]. \quad (3.9)$$

The first-order condition for location choice is

$$\frac{q'(x)}{q(x)} = \frac{-\tau_2(w\lambda, x)}{(1 - \alpha) [w\lambda - \tau(w\lambda, x) - p_c]}. \quad (3.10)$$

Given any rental function $q(\cdot)$, (3.10) determines the optimal location choice of a car-owner with ability $\lambda$, $x_c(\lambda)$. Similar to $x_b(\lambda)$, if $q(\cdot)$ is twice continuously differentiable, then $x_c(\lambda)$ exists and is continuously differentiable. If we define $\tilde{t}(w\lambda, x) \equiv \tau(w\lambda, x) + p_c$, the function $\tilde{t}(w\lambda, x)$ would share the same properties as $t(w\lambda, x)$, and the car-owners’ problem would then be isomorphic to the bus-users’ problem (P3.1). This implies that Lemma 3.1 can be applied on the car-owners’ problem.

If $x_c(\lambda)$ is the optimal location choice for car-owners, then it must yield positive values for consumption and land services, or

$$w\lambda - \tau[w\lambda, x_c(\lambda)] > p_c.$$
3.2.4 Car-ownership Decision

An agent with ability $\lambda$ will choose to own a car if and only if $V^c (\lambda) > V^b (\lambda)$. The car-ownership decision can be characterized by the function $\Omega (\lambda)$ where

$$
\Omega (\lambda) = \begin{cases} 
1 & \text{if } V^c (\lambda) > V^b (\lambda) \\
0 & \text{if } V^c (\lambda) \leq V^b (\lambda),
\end{cases}
$$

for $\lambda \in [\lambda_{\min}, \lambda_{\max}]$.

Substituting (3.8) and (3.9) into the utility function gives the value function of a car-owner,

$$
V^c (\lambda) = \tilde{\alpha} + \ln \{w \lambda - \tau [w \lambda, x_c (\lambda)] - p_c\} - (1 - \alpha) \ln q [x_c (\lambda)],
$$

where $\tilde{\alpha} \equiv \ln \alpha^\alpha (1 - \alpha)^{1-\alpha}$. Similarly, the value function of a bus-user is

$$
V^b (\lambda) = \tilde{\alpha} + \ln \{w - t [w \lambda, x_b (\lambda)]\} - (1 - \alpha) \ln q [x_b (\lambda)].
$$

Define a critical ability level $\overline{\lambda}$ such that $V^c (\overline{\lambda}) = V^b (\overline{\lambda})$. Then $\overline{\lambda}$ must satisfy

$$
\frac{w \overline{\lambda} - t [w \overline{\lambda}, x_c (\overline{\lambda})] - p_c}{w \lambda - t [w \lambda, x_b (\lambda)]} = \left\{ \frac{q [x_c (\overline{\lambda})]}{q [x_b (\overline{\lambda})]} \right\}^{1-\alpha}.
$$

This equation shows how location choices and car-ownership decisions are interdependent. If there does not exist any $\lambda$ in $[\lambda_{\min}, \lambda_{\max}]$ that satisfies (3.12), then the economy is said to have no car-owner.

Suppose condition (3.6) is satisfied for both $t(w \lambda, x)$ and $\tau(w \lambda, x)$, then $x_b (\lambda)$ and $x_c (\lambda)$ are both monotonically increasing. This implies that no bus-user will live further from the origin than $x_b (\overline{\lambda})$ and no car-owner will live closer to the
origin than \(x_c (\lambda)\). In equilibrium, \(x_c (\lambda) \leq x_b (\lambda)\) since, as discussed below, land rent will adjust so that agents are distributed continuously over the city. \(x_c (\lambda) < x_b (\lambda)\) implies that the location choices of car-owners overlap with those of bus-users. The following lemma shows that this is not possible.

**Lemma 3.2** *In equilibrium, no car-owner and bus-user will live at the same point. Hence, \(x_c (\lambda) = x_b (\lambda)\) holds.*

**Proof.** See Appendix. ■

### 3.2.5 Production

The aggregate output of final goods, \(Y\), and automobiles, \(A\), are produced using labor alone,

\[
Y = \eta L_Y,
\]

and

\[
A = z L_A,
\]

where \(L_Y\) and \(L_A\) denote the aggregate labor inputs devoted to the goods sector and the automobile sector, respectively. The variables \(\eta\) and \(z\) capture the TFP in the production of final goods and automobiles.

All markets are assumed to be competitive. Agents can choose to work in any one of the sectors. The market wage for an efficiency unit of labor is given by

\[
w = \eta = p_c z. \tag{3.13}
\]
3.3 Competitive Equilibrium

3.3.1 Definition

Boundary of City To simplify the analysis, the transportation costs functions are specified as

\[ t(w, x) = \tau_b w \lambda x + \gamma_b, \tag{3.14} \]

and

\[ \tau(w, x) = \tau_c w \lambda x + \gamma_c, \tag{3.15} \]

where \( \tau_b > \tau_c > 0 \) and \( \gamma_c + p_c > \gamma_b > 0 \). Under this specification, the critical ability, \( \bar{\lambda} \), and the corresponding location, \( x_c(\bar{\lambda}) = x_b(\bar{\lambda}) \equiv \bar{x} \), must satisfy

\[ \bar{x} = \frac{\gamma_c + p_c - \gamma_b}{(\tau_b - \tau_c) w \bar{\lambda}}. \tag{3.16} \]

The critical ability level, if it exists, is unique, with \( V^c(\lambda) > V^b(\lambda) \) if and only if \( \lambda > \bar{\lambda} \). Moreover, condition (3.6) stated in Lemma 3.1 is satisfied under this specification. Hence, \( x_b(\lambda) \) and \( x_c(\lambda) \) are both strictly increasing. The optimal location choice function for any agent is then given by

\[ x(\lambda) = x_c(\lambda) \Omega(\lambda) + x_b(\lambda) [1 - \Omega(\lambda)], \tag{3.17} \]

where \( x_b(\lambda) \) is implicitly defined by (3.5) and \( x_c(\lambda) \) is determined by (3.10). The function is continuous by the continuity of \( x_b(\lambda) \) and \( x_c(\lambda) \). Strict monotonicity of \( x_b(\lambda) \) and \( x_c(\lambda) \) implies that \( x(\lambda) \) is also strictly monotonic. Thus, no one will reside beyond \( x_c(\lambda_{\text{max}}) \), where \( \lambda_{\text{max}} \) is the maximum ability in the ability
distribution (assumed to be finite). Define \( \tilde{x} \equiv x_c(\lambda_{\text{max}}) \), the size of the city is

\[ C = [0, \tilde{x}] . \]

In equilibrium, every point in \( C \) must be occupied. Otherwise, any rational landowner would lower the rent at the empty point so as to induce someone to move in. Hence, \( x(\lambda) \) should be continuous over the range of abilities, \([\lambda_{\text{min}}, \lambda_{\text{max}}]\).

Land beyond the boundary of the city is used for agriculture. Let \( q_A \) denote the agricultural rent. If \( q(\tilde{x}) < q_A \), then any rational land-owner would choose to rent out the land at \( \tilde{x} \) to agricultural users. If \( q(\tilde{x}) > q_A \), then households at the boundary would be strictly better off by moving slightly further away from the city center. The reason is, if the movement is sufficiently small, then transportation costs would only go up marginally but would be compensated by a significant reduction in rent, \([q(\tilde{x}) - q_A]\). Thus, in equilibrium,

\[ q(\tilde{x}) = q_A . \quad (3.18) \]

**Population Density** Let \( f(\lambda) \) be the density function governing the ability distribution. To derive the population density function, consider an ability \( \lambda \) and a small neighborhood of length \( d\lambda \) around it. The fraction of population within this neighborhood is \( f(\lambda) d\lambda \). Since \( x(\lambda) \) is strictly monotonic and hence one-to-one, one can find a neighborhood \( dx \) around \( x = x(\lambda) \) such that all agents with ability \( \lambda \in d\lambda \) are located in this interval. Population density at \( x \) is, hence, given by

\[ \mu(x) = \frac{f(\lambda) d\lambda}{dx} \]
which implies
\[ \mu [x(\lambda)] x'(\lambda) = f(\lambda), \] (3.19)
for any \( \lambda \in [\lambda_{\min}, \lambda_{\max}] \). Alternatively, (3.19) can be derived using the transformation of variable technique. The equilibrium population density function over \( C \) is then defined as
\[ \mu(x) = \begin{cases} 
\mu [x_b(\lambda)] & \text{if } x = x_b(\lambda) \in [0, \bar{x}], \\
\mu [x_c(\lambda)] & \text{if } x = x_c(\lambda) > \bar{x}.
\end{cases} \] (3.20)

**Market Clearing Conditions** In equilibrium, the land demand function for any type-\( \lambda \) agent is given by
\[ \tilde{l}(\lambda) = \begin{cases} 
l_b [x_b(\lambda), \lambda] & \text{for } \lambda \leq \overline{\lambda}, \\
l_c [x_c(\lambda), \lambda] & \text{for } \lambda > \overline{\lambda}.
\end{cases} \]
Consider any subinterval \([\lambda_1, \lambda_2] \subseteq [\lambda_{\min}, \lambda_{\max}]\). Since \( x(\lambda) \) is strictly increasing, there exists a unique interval \( C' = [x(\lambda_1), x(\lambda_2)] \) in \( C \) that contains all the agents in \([\lambda_1, \lambda_2]\). Land markets in \( C' \) clear when the total demand for land equals the total supply
\[ \int_{\lambda_1}^{\lambda_2} \tilde{l}(\lambda) f(\lambda) d\lambda = \int_{x(\lambda_1)}^{x(\lambda_2)} dx = \int_{\lambda_1}^{\lambda_2} x'(\lambda) d\lambda. \] (3.21)
Since (3.21) has to hold for all \([\lambda_1, \lambda_2]\), the land market clearing condition at any \( x = x(\lambda) \) can be restated as
\[ \tilde{l}(\lambda) f(\lambda) = x'(\lambda), \quad \text{for } \lambda \in [\lambda_{\min}, \lambda_{\max}]. \] (3.22)
Agents supply their non-commuting hours to market work so labor supply depends on their transportation mode and location choices. The labor market clears if the following holds:

\[ L_A + L_Y = \int_{\lambda_{\min}}^{\lambda_{\max}} \{1 - \tau_b x_b(\lambda) [1 - \Omega(\lambda)] - \tau_c x_c(\lambda) \Omega(\lambda)\} \lambda dF(\lambda). \] (3.23)

Aggregate demand for automobiles is given by the fraction of population that choose to own a car. Hence, the auto market clears when

\[ A = \int_{\lambda_{\min}}^{\lambda_{max}} dF(\lambda). \] (3.24)

Final goods produced in this economy, net of those dissipated in commuting, are available for consumption. Aggregate demand for consumption is the sum of the demand by agents and that by the landowner. Since all the rents collected by the landowner are spent on consumption, the demand for consumption by the landowner is

\[ Q = \int_{\lambda_{\min}}^{\lambda_{\max}} \{q [x_b(\lambda)] l_b(\lambda) [1 - \Omega(\lambda)] + q [x_c(\lambda)] l_c(\lambda) \Omega(\lambda)\} dF(\lambda). \]

Total demand for consumption goods by the agents is

\[ C = \int_{\lambda_{\min}}^{\lambda_{\max}} \{c_b(\lambda) [1 - \Omega(\lambda)] + c_c(\lambda) \Omega(\lambda)\} dF(\lambda). \]

The final goods market clear when

\[ Y = \left[ \gamma_b \int_{\lambda_{\min}}^{\lambda_{\max}} dF(\lambda) + \gamma_c \int_{\lambda_{\min}}^{\lambda_{\max}} dF(\lambda) \right] = Q + C. \] (3.25)
Define the equilibrium of this economy as follows:

**Definition**  Given a distribution of abilities, \( F(\lambda) \), an equilibrium of this economy consists of a set of decision rules for car-owners, \( \{c_c(\lambda), l_c(\lambda), x_c(\lambda)\} \), a set of decision rules for bus-users, \( \{c_b(\lambda), l_b(\lambda), x_b(\lambda)\} \), labor inputs, \( \{L_A, L_Y\} \), a car-ownership decision rule, \( \Omega(\lambda) \), a population density function, \( \mu(x) \), a critical ability level, \( \bar{\lambda} \), a rental function, \( q(\cdot) \), and prices, \( \{p_c, w\} \) such that

1. Given \( q(\cdot), p_c, \) and \( w \), \( \{c_b(\lambda), l_b(\lambda), x_b(\lambda)\} \) solves (P3.1).
2. Given \( q(\cdot), p_c, \) and \( w \), \( \{c_c(\lambda), l_c(\lambda), x_c(\lambda)\} \) solves (P3.2).
3. The prices, \( p_c \) and \( w \), satisfy (3.13).
4. The car-ownership decision rule, \( \Omega(\lambda) \), is given by (3.11).
5. The population density function, \( \mu(x) \), defined by (3.20), satisfies
   \[
   \int_c \mu(x) \, dx = 1.
   \]
6. The rent at the boundary of the city equals the agricultural rent, or (3.18) holds.
7. All markets clear:
   (a) The land market at every \( x \) clears, or (3.22) hold.
   (b) The auto market clears or (3.24) holds.
   (c) The final good market clears or (3.25) holds.
Equilibrium Rent  In equilibrium, the rent function must be continuous over the range of the city. Given a continuous distribution of abilities and a continuous transportation cost function, it is immediate to see that $q(x)$ is continuous over the region where there are bus-users (or car-owners) alone. Hence $q(x)$ must be continuous over $C \setminus \{\pi\}$. Suppose it was discontinuous at $\pi$, and

$$q(\bar{x}) > \lim_{x \to \bar{x}^+} q(x).$$

Consider an agent with the critical ability, $\bar{\lambda}$. This agent lives at $\bar{x}$. Given the gap in rent at $\pi$, he can benefit by moving slightly further out to a location $\bar{x} + \varepsilon$. Since $\varepsilon$ can be made arbitrarily small, a new location can always be found such that the agent’s additional transportation costs are less than his discrete gain from savings in rent. This creates an incentive to move and hence cannot be an equilibrium. By a similar argument, one can rule out the case with $\lim_{x \to \bar{x}^-} q(x) > q(\bar{x})$. Hence, $q(x)$ is continuous over $C$.

3.3.2 Characterization

An equilibrium is made up of three parts: the bus-user’s problem, the car-owner’s problem and a critical ability level $\bar{\lambda}$ that connects the two. The bus-user’s problem is characterized by

$$q'[x_b(\lambda)] = \frac{-w\lambda \tau_b}{q[x_b(\lambda)]} \left(1 - \alpha\right) \left\{w\lambda[1 - \tau_b x_b(\lambda)] - \gamma_b\right\},$$

and

$$x'_b(\lambda) = \frac{(1 - \alpha) \left\{w\lambda[1 - \tau_b x_b(\lambda)] - \gamma_b\right\} f(\lambda)}{q[x_b(\lambda)]}.$$
for $\lambda \in [\lambda_{\text{min}}, \lambda]$, and the boundary conditions:

\[
\begin{align*}
    x_b(\lambda_{\text{min}}) &= 0, \\
    x_b(\bar{\lambda}) &= \bar{x}.
\end{align*}
\] (3.28)

Define the composite functions $q_b(\lambda) \equiv q[x_b(\lambda)]$ and $\mu_b(\lambda) = \mu[x_b(\lambda)]$. Combining (3.26) and (3.27) gives

\[
q'_b(\lambda) = -w\lambda \tau_b f(\lambda).
\] (3.29)

Notice that the above expression is independent of $q_b(\lambda)$ and $x_b(\lambda)$. This happens because (i) land supply is fixed and constant at each location and (ii) given $w\lambda$, the marginal costs of commuting, $t_2(w\lambda, x)$, are identical for all locations. Integrating both sides of (3.29) gives

\[
q_b(\lambda) = \psi - \tau_b w \int_{\lambda_{\text{min}}}^{\lambda} u f(u) du,
\] (3.30)

for $\lambda \in [\lambda_{\text{min}}, \lambda]$, where $\psi$ is the integration constant.

Similarly, the car-owner’s problem is characterized by

\[
x'_c(\lambda) = \frac{(1 - \alpha)\{w\lambda [1 - \tau_c x_c(\lambda)] - (\gamma_c + p_c)\} f(\lambda)}{q_c(\lambda)},
\] (3.31)

and

\[
q_c(\lambda) = q_A + \tau_c w \int_{\lambda}^{\lambda_{\text{max}}} u f(u) du,
\] (3.32)

for $\lambda \in [\bar{\lambda}, \lambda_{\text{max}}]$, and the boundary condition:

\[
x_c(\bar{\lambda}) = \bar{x}.
\]
Finally, the critical ability level $\bar{\lambda}$ and the corresponding location $\bar{x}$ must satisfy (3.16). The algorithm employed in the computation exercise is outlined in the Appendix.

### 3.4 Quantitative Analysis

Can the model account for the decreasing population density gradients and rising level of car-ownership? To see how well the model can do at explaining suburbanization and car-ownership consider the following experiment. Compute the model for a series of steady states. The steady states represent an average U.S. city during the decennial years from 1910 to 1970. Then, compare the model’s prediction for the density gradients and car-ownership with the mean density gradients of forty-one U.S. cities, presented in Table 3.3, and the percentages of car-owners in the U.S., approximated by the number of registered vehicles per person age twenty to sixty-four, given in Table 3.4.

Before the steady states can be computed, a parametrization must be imposed on the model. This is done through a combination of calibration and estimation. The strategy behind the calibration is as follows. First, parameters that can be pinned down from the U.S. data will be calibrated accordingly. Those that cannot will be chosen in order to minimize the “distance” between the model’s outcome and the data. Under this calibration strategy, the model is given its “best shot” to match the data on suburbanization and car-ownership. In order to gauge the success of the model, its predictions on city sizes and transportation costs are then compared to the observed data.

---

15 Note that the “estimation” procedure used here is not based on any statistical model and hence is not the basis for any inference.
3.4.1 Parameters to Set

Let one period of the model equal five years, approximately the average median age of passenger cars during the period 1950-70.\footnote{The average median age of passenger cars in the U.S. during the period 1950-70 was 5.1 years. Source: Ward’s Motor Vehicle Facts & Figures (1999).} The parameters which need to be set are:

**Preferences** The parameter, $\alpha$, measures the relative weight on consumption in the utility function. Set $\alpha$ to 0.89 using data on consumer expenditures from Lebergott (1996).

**Technology** The parameters, $\eta_t$ and $z_t$, which capture the TFP in production of final goods and automobiles, are chosen such that $w_t$ equals the mean income levels for each steady state, and $p_{c,t}$ at each steady state matches the data on the price of a new car.\footnote{Subscript ‘$t$’ implies that the parameter is not assumed to be constant across the steady states.} The mean income level and automobile prices are based on the data on wages and car prices presented in Section 3.1.2.

**Abilities** The distribution of abilities is approximated by a doubly-truncated lognormal distribution. The standard deviation is calibrated so that the distribution of income ($w\lambda$) matches the distribution of income in the U.S.\footnote{According to Gottschalk and Smeeding (1997), Table 3, the adjusted disposable personal income of a household at the 80th percentile is 2.7 times higher than one at the 20th percentile. This implies that, if the ln of income is normally distributed, then the standard deviation is 0.59.} Given the mean and standard deviation, the truncation points, $\lambda_{\text{min}}$ and $\lambda_{\text{max}}$, are chosen so as to encompass 95% of the area of the underlying non-truncated distribution, omitting 2.5% from each side.
Agricultural Rent  The rental rate of agricultural land, \( q_{A,t} \), is set to the rent that would have to be paid for an average single-family-sized lot of farmland at each date.\(^{19}\)

Transportation Costs  The parameters related to transportation costs, namely, the fixed and time costs for taking a bus, \( \gamma_{b,t} \) and \( \tau_{b,t} \), and those for commuting by car, \( \gamma_{c,t} \) and \( \tau_{c,t} \), are the most difficult to determine. These parameters must capture all the costs associated with commuting. For car-owners these include all costs associated with operating a car, plus the time costs of commuting. For bus-users, these include bus fares and time costs. The time cost of taking the bus, as opposed to commuting by car, includes not only the time spent on the bus but also all the inconveniences associated with using public transportation. Since it is difficult to measure all of these costs and their changes over time, a combination of calibration and estimation is used to determine these parameters.

(i) Automobile: The time cost of commuting by car depends on, among other things, the quantity and quality of roads and highways. Throughout the period 1910-1970 the U.S. government consistently invested in roads and highways. The stock of U.S. highways and roads per capita rose at an average rate of 2% during the period 1925-1994, and at an average rate of 2.3% over the period 1925-70. Figure 3.9 plots the highway stock per capita over the period 1925-1994.\(^{20}\) Assume that the time cost of commuting is

---

\(^{19}\)The rental rate for an average single-family-sized lot of farmland is computed by dividing the gross rent paid for an acre of farmland, taken from the United States Department of Agriculture Economic Research Service and is based on various sources including the Census of Agriculture, the Farm Costs and Returns Survey, and the Farm Finance Survey, by the number of average-sized lots in an acre. To compute the number of average-sized lots in an acre, data on the average lot size for a single-family home is taken from the National Association of Realtors, available on the web at <http://www.realtor.org/SG3.nsf/files/landuse.pdf/$FILE/landuse.pdf>.

\(^{20}\)Source: Fixed Reproducible Tangible Wealth in the United States.
inversely related to the highway and road stock per capita. Then $\tau_{c,t}$ can be approximated by the following:

$$\tau_{c,t} = \Phi h_t^{-\kappa},$$

where $h_t$ is the capital stock of highways and streets per capita at time $t$. The parameters $\Phi$ and $\kappa$ are chosen through the estimation procedure discussed in Section 3.4.2. The value for the highway and street stock for the years 1910 and 1920 are obtained by extending the trend line out to these dates. The fixed cost, $\gamma_{c,t}$ is calibrated using data on consumer car-related expenditures from Lebergott (1996). The data include expenditures on tires, gas, oil, automobile repairs, automobile insurance, and tolls. The expenditure per registered vehicle is computed. Then $\gamma_{c,t}$ is set to 30%, the percentage of total miles that are driven to or from work, of period $t$’s value.\(^{21}\)

(ii) **Bus**: The parameters $\tau_{b,t}$ and $\gamma_{b,t}$ are difficult to calibrate since they must capture the costs associated with the inconveniences of using public transportation. Hence these parameters are derived through an estimation procedure. Similar to the time cost of commuting by car, it is intuitive that the time cost of travelling by bus is inversely related to the quantity and quality of public transit equipment and structures per capita. Figure 3.10 shows the gross and net capital stock per capita of equipment and structures for intercity and local passenger transit for the period 1947-70.\(^{22}\) Both the

---


\(^{22}\)Source: Fixed Reproducible Tangible Wealth in the United States.
gross and net stock have been steadily declining since 1947 at an annual rate of approximately 6% over the period 1947-70. Figure 3.11 shows the trend in transit ridership per capita over the period 1910-70. According to the graph, transit ridership rates peaked sometime during the 1920’s. Table 3.5 shows the trend in net capital expenditures on transit equipment and structures for the period 1890-1950. The table shows that disinvestment in public transportation started during the 1920’s. It appears that the capital stock in public transit has been declining since, at least, 1930. As the rate of transit ridership and the stock of transit capital declines, the cost associated with using public transportation should increase. With less riders it is logical to assume that buses, subways, etc. ran less frequently and that their routes covered smaller portions of cities. The declining capital stock implies that the number of buses, etc. actually declined. Together these factors indicate that the cost of commuting by public transit should be rising since, at least, 1930 due to the increasing inconvenience and unavailability bus-users faced. As for the time cost of commuting by car, the time cost of commuting by bus is assumed to be inversely related to the capital stock in public transit. But there is no data on this stock for years preceding 1947. So, instead, consider the following formulation:

\[
\tau_{b,t} = \begin{cases} 
\tau_{b,0}(g_{\tau_b})^{(t-1910) / 10}, & \text{for } 1910 \leq t \leq 1930, \\
\Omega p_t^{-\rho}, & \text{for } t \geq 1930.
\end{cases}
\]

Here, \(\tau_{b,t}\) is a function of the capital stock of public transit, \(p_t\), for the years 1930 to 1970 and is assumed to grow at a constant rate for the years 1910 to

\(^{23}\text{Source: Jones (1985).}\)
1930. For the years 1930-47, the capital stock is computed by extending the trend line out to these dates. The trend is not extended to the years 1910-30 because it is questionable whether the capital stock was declining during these early years. Hence $\tau_{b,t}$ at each time $t$ will be determined by choosing $\Omega, \rho, \tau_{b,0}$, and $g_{\tau_b}$ through the estimation procedure described below subject to the constraint:

$$\tau_{b,0}(g_{\tau_b})^{(\frac{1930-1910}{10})} = \Omega p_{1930}^{-\rho},$$

or, in the year 1930, the time cost implied by assuming a constant growth rate of $\tau_{b,t}$ must be equal to the one implied by assuming that $\tau_{b,t}$ is inversely related to the public transit stock per capita. The fixed cost of commuting by bus, $\gamma_{b,t}$, should capture costs such as bus fares and the inconvenience of taking the bus relative to commuting with a car, but data on average bus fares and other costs of taking the bus in the U.S. for this time period are not available. Hence this parameter is derived through the estimation procedure. To enforce some structure on this cost, it is assumed to grow at some constant rate over the period 1910-1970. Let $g_{\gamma_b}$ be the growth factor of $\gamma_b$, and $\gamma_{b,0}$ be the initial value in 1910. Then

$$\gamma_{b,t} = \gamma_{b,0}g_{\gamma_b}^{(\frac{t-1910}{10})}.$$

The parameters $\gamma_{b,0}$ and $g_{\gamma_b}$ are chosen through the estimation procedure described below.

### 3.4.2 Estimation

Denote by $v_t$ the percentage of car-owners in the U.S. at date $t$. This can be approximated by the number of registered vehicles per person at age twenty to
sixty-four, given in Table 3.4. Likewise, let \( d_t \) be the average density gradient for an American city. The values of \( d_t \) are reported in Table 3.3.

Define the following vector of unknown parameters:

\[
\theta = (\Phi, \kappa, \tau_{b,0}, g_{\tau_b}, \Omega, \rho, \gamma_{b,0}, g_{\gamma_b}).
\]

Given \( \theta \), an ability distribution \( F_t(\cdot) \), and a price for cars, \( p_{c,t} \), the model’s prediction for the percentage of car-owners at date \( t \) is denoted by

\[
V_t(\theta; F_t(\cdot), p_{c,t}).
\]

Likewise, the model’s predicted density gradient is

\[
D_t(\theta; F_t(\cdot), p_{c,t}).
\]

To compute \( D_t(\theta; F_t(\cdot), p_{c,t}) \), the density is calculated at 1000 points from zero to the end of the city. The density gradients are then computed by regressing the sample on an exponential function. In all the experiments below, the R-squared’s are always above 0.80.

The exercise, now, consists of two steps: First, \( \theta \), is chosen to minimize the sum of the deviations between the model’s output and the U.S. economy at a particular set of steady states corresponding to the decennial years from 1910 to 1970. Formally,

\[
\hat{\theta} = \arg \min_{\theta} \sum_{t=1910}^{1970} \left\{ \frac{1}{2} [v_t - V_t(\theta; F_t(\cdot), p_{c,t})]^2 + \frac{1}{2} [d_t - D_t(\theta; F_t(\cdot), p_{c,t})]^2 \right\}.
\]

Second, the model’s predictions, \( V_t(\hat{\theta}; F_t(\cdot), p_{c,t}) \) and \( D_t(\hat{\theta}; F_t(\cdot), p_{c,t}) \), for \( t = \)
3.5 Findings

3.5.1 Baseline Results

The baseline parametrization is presented in Table 3.6. The table also displays statistics on the costs of commuting, implied by the parameters. The statistics given are the fixed, variable, and total costs of commuting for a bus-user at the average distance from the center and a car-owner at the average distance from the center. These will be discussed in Section 3.5.2.

The results are presented in Table 3.7. The table shows the model’s predictions for the percentage of car-owners, the population density gradient, and the change in the population density gradient over the period 1910-1970. In addition, the table provides the distance at which the furthest bus-user lives (End of Bus-Users) and that at which the furthest car-owner lives (Boundary of City) in each steady state.

3.5.2 Discussions

In this section the baseline model and its ability to match the data is assessed. To begin consider the transportation costs, specifically, consider those associated with taking public transit, or “riding the bus”. The fixed and time cost of riding the bus are rising over the period 1910-1970. Since bus-users move closer and closer to the origin and bus-users who switch to car-ownership are always those with the highest incomes, the average variable cost of travelling by bus does not rise over the period, it declines during 1910-40, then rises between 1940-50 and

\[ \hat{\theta} \]

A similar calibration procedure is used in Andolfatto and MacDonald (1998).
then declines during 1950-70. The average total cost declines at an annual rate of 3.7% during 1910-30 and then rises at an annual rate of 2.9% between 1930-70. The minimization procedure generated results that are consistent with the assumption that the cost of commuting by public transit is inversely related to the capital stock of public transit equipment and structures.

The fixed costs of commuting by car was already calibrated to match the data. The time cost is declining, as was expected given the rising capital stock of highways and roads. Since the model predicts zero car-owners in 1910, it is not possible to compute the variable and fixed costs for this year. The average variable cost falls between 1920 and 1930 but rises from 1930 to 1960 then falls slightly in 1970. The average total cost follows a similar trend to the variable cost. This is because average variable and total costs depend on both the income levels and distance from the center. For the car-owners both of these variables are rising over the period. Taken into account the rising fixed costs, the fact that average total costs for car-owners are increasing does not seem unreasonable.

Now, let’s approach these costs at a different, more quantitative angle. There is significantly more data available for later years. Therefore, for the moment, consider comparing the transportation costs associated with commuting by public transit and by automobile in the steady state corresponding to the year 1970 with the data. The goal is to determine if the magnitude of these costs are reasonable. Focus on the average total costs. In 1970 this cost for a bus-user is $9.25 per day. Based on the 1975 Census, the mean travel time to work in 1975 by those who take public transportation is 40.1 minutes. This implies a total daily travel time of 80.2 minutes. The average hourly wage of a bus-user in 1970 is $6.15.\textsuperscript{25} This

\textsuperscript{25}This is the wage of the average bus-user and is not the same as the wage of the bus-user who is at the average distance which is given in Table 3.6.
implies the dollar cost of time spent commuting is $8.22 a day. Hence, considering the additional cost of bus-fare, $9.25 a day is a very reasonable value. For the average car-owner the total cost of commuting in 1970 is $29.88. Again from the 1980 Census, the mean travel time to work in 1980 of people who commute by car is 21.6 minutes. This implies that the total commute time is 43.2 minutes. The average hourly wage of a car-owner in 1970 is $23.22. Hence the value of time spent commuting by car is $16.72 a day. The distance from the center of this average car-owner is 7.77 miles.\(^{26}\) The total cost of operating a vehicle in 1975 was approximately 58.59 cents/mile.\(^{27}\) Thus, the total “car-related” cost for the average car-owner in the model is $4.55 a day. Summing the car-owner’s costs suggests that his total costs of commuting by car are $21.27 a day. So a total cost of $29.88 a day resulting from the minimization is very reasonable.

How are the results? Qualitatively, the model is able to predict a persistent increase in car-ownership accompanied by increased suburbanization. The latter is evident from the decline in the population density gradient and the expansion of the city as shown in Table 3.7. The model is able to match the data on car-ownership for earlier years but has some difficulty reaching the level of car-ownership in later years. One reason for this is that the model abstracts from other uses of cars. In 1969, about 32% of personal trips by car involved work travel, this number decreases to 23.8% by 1995.\(^{28}\) Meanwhile, vehicle trips that involve shopping and, social and recreational activities experience an opposite trend. By 1995, about 40% of personal trips by car are for these purposes. By

\(^{26}\) Again, this is not the same average as in Table 3.6.

\(^{27}\) Source: American Automobile Association. (1993) “Your Driving Costs." the total cost includes fixed costs—depreciation, insurance, finance charge, and license fee, and variable costs—gas, oil, maintenance, and tires.

\(^{28}\) Source: U.S. Federal Highway Administration, Summary of Travel Trends, 1995 National Personal Transportation Survey.
including additional uses of car into the model, the demand for automobiles could be raised. In terms of the population density gradient, the model is able to match the trend but underestimates the density gradient for later years.

The average distance between home and work in the U.S. was 6.4 miles in 1955 and 9.4 miles in 1969.\textsuperscript{29} In the model the average distance of a car-owner is 6.19 miles in 1950, 6.95 miles in 1960, and 8.13 miles in 1970. It seems that the distances in the model are within a reasonable magnitude.

### 3.5.3 Counterfactual Experiments

Table 3.8 provides the results of a series of counterfactual experiments. Each experiment consists of shutting down one or some combination of the factors that lead to suburbanization and car-ownership: rising real wages, falling prices of cars, declining cost of commuting by car, and rising costs of commuting via public transportation. The counterfactual experiments can be used to assess the role that various factors play in generating the increasing trend in car-ownership and decreasing density gradient.

The baseline model is able to account for 86% of the car-ownership in 1970 and 77% of the difference between the density gradient in 1910 and that in 1970. The first three experiments focus on the impact of both rising real wages and falling prices of automobiles. When both prices and wages remain fixed at their 1910 values (Experiment 3a), the model accounts for less than 1% of car-ownership. The density gradient in this case actually increases. This is due to two factors: First, 1910 wages and car prices mean that, in all time periods, most agents are unable to afford a car. Second, the cost of riding the bus is rising over time.

Hence the agents move towards the origin instead of away from it. Removing the rising costs of public transit (Experiment 3b) removes the remaining incentive to purchase a car. Hence, in this case, no one chooses to be a car-own-er and there is no suburbanization. If only the car price remains at the 1910 value (Experiment 4a) then the model can account for 51% of car-ownership in 1970, and 60% of the change in the density gradient. A comparison between Experiments 3a and 4a illustrates how rising income levels can promote car-ownership and urban sprawl. The positive relationship between car-ownership and suburbanization is weakened, though, when the rising cost of using public transit is removed as in Experiment 4b. In the case, the model only accounts for 13% of car-ownership and 65% of suburbanization. This suggests that in the presence of an efficient public transport system, it is possible to obtain a suburbanization trend with a small percentage of car-ownership. When the wage distribution remains as in 1910 (Experiment 5a), then 21% of car-ownership in 1970 is accounted for and the density gradient rises, but removing the rising time costs of public transit (Experiment 5b) reduces the incentive to switch to car, resulting in no car-owners and again less centralization.

The remaining three experiments examine the role of changing commuting costs. Experiment 6 in the table shows how the results change when the highway and street stock per capita is kept fixed at its 1910 value. Since the stock is not rising in this experiment, the time cost of commuting by car remains constant and the cost to car-owners of spreading out is high. The percentage of car-owners still rises over time but only reaches 74% of the percentage in the baseline model. Despite the significant percentage of car-owners, 54% in 1970, there is no suburbanization trend. In fact agents become more concentrated around the city center. This result occurs because the cost of commuting by bus is rising, pushing bus-users closer to the center, while the cost of commuting by car does not
fall, removing the incentive for car-owners to spread further out. In experiment 7 the time and fixed costs of commuting by bus are kept constant at their 1910 values. Notice that, without a rising cost associated with taking the bus, the percentage of agents who switch to car-ownership is much lower than in the baseline results. Under the parametrization of experiment 7 only 31% of car-ownership can be explained, but the suburbanization trend is actually stronger than the trend predicted by the baseline model. These results are similar to those obtained in Experiment 4b: with falling car prices, more people switch to commuting by car and this encourages urban sprawl.

Finally, in experiment 8, the time cost of travelling by car is kept at its 1910 value and, in addition, the time and fixed cost of commuting by bus are kept at their 1910 values. Under this parametrization the bus is as attractive as possible, relative to the car, without violating the assumption that the car is a form of transportation which has a lower time cost than the bus.30 Here no agent becomes a car-owner and, in addition, despite the rising real wages, no suburbanization trend is observed. To summarize, the experiments suggest the following:

1. The suburbanization trend is caused by a combination of rising real wages and the diffusion of a transportation technology that becomes more affordable and improves in efficiency over time. In other words, rising real wages and falling prices of automobiles alone are not enough to generate a suburbanization trend (Experiments 6 and 8), nor is a declining time cost associated with automobiles alone able to generate a suburbanization trend (Experiment 3).

---

30In this model, decreasing the time cost of commuting by bus over time starting from the 1910 value would violate the definition of the bus as a form of transportation which has a higher time cost relative to the car.
2. No suburbanization trend can be driven by bus-users alone (Experiments 3b, 5b, and 8). Yet, in the presence of an efficient public transport system, there can be a suburbanization trend with a much smaller percentage of car-owners than what is observed in the data (Experiment 7).

3. A high percentage of car-owners does not necessarily imply a suburbanization trend. If public transportation becomes increasingly unattractive due to rising time costs and there is no technological progress in automobile use then there can be a high percentage of car-ownership but no suburbanization. (Experiment 6).

3.6 **Concluding Remarks**

A general equilibrium model of car-ownership and location choice is constructed. An agent, in the model, can choose his residential location and decide whether or not to own a car. Under the given specification, it is shown that wealthy agents (those with high abilities) tend to own a car and live further away from the city center, while poor agents tend to travel by bus and stay close to the city center. The model is then calibrated using U.S. data. It is able to predict the rising trends in both car-ownership and suburbanization. Despite the fact that a highly stylized framework is used, the model is able to explain 86% of car-ownership in 1970 and roughly 77% of suburbanization between 1910 and 1970 as reported in the data.

There are some important features of car-ownership that are not addressed by the model. The model abstracts from other uses of cars. By including additional uses of cars into the model, the demand for automobiles could be raised. This might help to improve the model’s prediction on car-ownership. In the U.S., mul-
tiple car-ownership increases significantly during the latter half of the twentieth century. For instance, less than 10% of American households owned more than one car in the early 1950s, yet by 1995, 60% of them did. The same period also recorded a rapid increase in female labor-force participation rate. Data suggest that female participation and car-ownership decisions might be closely related. Two-income households are more likely to own a car than households in which only the man works. Moreover, the proportion of families with more than one car is rising with the wife’s earnings. To account for these facts, the model can be extended to include female workers. By having an additional worker in the household, there is an extra source of income and the need for a second car.

Given the static nature of the model, car-ownership decision is based solely on current prices and income level. In the presence of declining car price and rising income, agents might delay their car-ownership decisions. To explore the consequences of this type of forward-looking behavior, a dynamic framework is needed.

Future versions of the model would also benefit from modeling the development of roads and highways more explicitly. Clearly, the role that declining transportation costs for car-owners and rising transportation costs for bus-users play is a significant one in determining the relationship between car-ownership and suburbanization. Potentially adding an urban developer or a government to the model who creates roads and/or public transportation would be helpful.
3.7 Appendix

Proofs

Proof of Lemma 3.1. To derive the slope of $x_b(\lambda)$, first totally differentiate equation (3.5)

$$ \frac{q'(x)}{q(x)} = -\frac{-t_2(w, x)}{(1-\alpha)[w - t(w, x)]}$$

with respect to $x$ and $\lambda$. This gives

$$ \left[ \frac{q''}{q} - \left( \frac{q'}{q} \right)^2 \right] dx = \frac{-[t_{22}dx + t_{12}wd\lambda]}{(1-\alpha)[w - t(w, x)]} - \frac{(-t_2)(1-t_1)wd\lambda + (-t_2)^2 dx}{(1-\alpha)[w - t(w, x)]^2}.$$  

(3.33)

Consider the LHS of the above expression,

$$ LHS = \left\{ \frac{-t_{22}}{(1-\alpha)[w - t(w, x)]} - \frac{(-t_2)^2}{(1-\alpha)[w - t(w, x)]^2} \right\} dx + \left[ \frac{wt_{12}}{t_2} - \frac{w(1-t_1)}{w - t(w, x)} \right] \frac{(-t_2)d\lambda}{(1-\alpha)[w - t(w, x)]} $$

$$ = \left[ \frac{t_{22}q'}{t_2 q} - (1-\alpha) \left( \frac{q'}{q} \right)^2 \right] dx + \left[ \frac{wt_{12}}{t_2} - \frac{w(1-t_1)}{w - t(w, x)} \right] \left( \frac{q'}{q} \right) d\lambda. $$

The second equality is obtained by using (3.5). Equality (3.33) can then be simplified as

$$ \left[ \frac{q''}{q} - \alpha \left( \frac{q'}{q} \right)^2 - \frac{t_{22}q'}{t_2 q} \right] dx = \left[ \frac{w\lambda t_{12}}{t_2} - \frac{w\lambda(1-t_1)}{w - t(w, x)} \right] \left( \frac{q'}{q} \right) d\lambda. $$

The desirable results can be obtained if the following inequality holds

$$ \left[ \frac{q''}{q} - \alpha \left( \frac{q'}{q} \right)^2 - \frac{t_{22}q'}{t_2 q} \right] > 0. $$  

(3.34)
The reason is, since \( q(x) \) is decreasing, \( x_b(\lambda) \) is monotonically increasing if and only if

\[
\frac{w\lambda t_{12}}{t_2} < \frac{w\lambda (1 - t_1)}{w\lambda - t(w\lambda, x)}
\]

for all positive \( w\lambda \) and \( x \). The RHS of the above inequality is the percentage increase in net income when labor income, \( w\lambda \), increase by one percentage. Given the log utility, this coincides with the income elasticity of housing demand. Hence, what remains is to show that condition (3.34) holds. It turns out that (3.34) is a sufficient condition to ensure that \( x_b(\lambda) \) is an interior solution for the maximization problem.

Given (3.2) and (3.3), the optimization problem (P1) is equivalent to

\[
\max_{x \in [0, \phi]} \{ U(x) = \ln [w\lambda - t(w\lambda, x)] - (1 - \alpha) \ln q(x) \}.
\]

The first-order and second-order derivatives are given by

\[
U_x = \frac{-t_2}{w\lambda - t(w\lambda, x)} - (1 - \alpha) \frac{q'(x)}{q(x)},
\]

\[
U_{xx} = \frac{-t_{22}}{w\lambda - t(w\lambda, x)} - \frac{(-t_2)^2}{[w\lambda - t(w\lambda, x)]^2} - (1 - \alpha) \left[ \frac{q''}{q} - \left( \frac{q'}{q} \right)^2 \right].
\]

Given that \( U_x = 0 \), \( U_{xx} < 0 \) if and only if

\[
(1 - \alpha) \left[ \frac{q''}{q} - \left( \frac{q'}{q} \right)^2 \right] > \frac{-t_{22}}{w\lambda - t(w\lambda, x)} - \frac{(-t_2)^2}{[w\lambda - t(w\lambda, x)]^2}
\]

\[
\Leftrightarrow (1 - \alpha) \left[ \frac{q''}{q} - \left( \frac{q'}{q} \right)^2 \right] > \frac{-t_{22}}{w\lambda - t(w\lambda, x)} - (1 - \alpha)^2 \left( \frac{q'}{q} \right)^2
\]
\[
\iff \left[ \frac{q''}{q} - \alpha \left( \frac{q'}{q} \right)^2 \right] > \frac{-t_{22}}{(1 - \alpha) [w\lambda - t(\lambda, x)]}.
\]

**Proof of Lemma 3.2.** Let \( W^c(x; \lambda) \) denote the value function of a car-owner with ability \( \lambda \) who locates at \( x \). By definition,

\[
V^c(\lambda) \equiv W^c[x^c(\lambda); \lambda] \geq W^c(x; \lambda) \quad \text{for all } x \geq 0.
\]

Similarly, define \( W^b(x; \lambda) \) for bus-users. Then,

\[
V^b(\lambda) \equiv W^b[x^b(\lambda); \lambda] \geq W^b(x; \lambda) \quad \text{for all } x \geq 0.
\]

Suppose \( \lambda^* \) exists so that \( V^c(\lambda^*) = V^b(\lambda^*) \). This implies

\[
V^b(\lambda^*) \geq W^c[x^b(\lambda^*); \lambda^*] \quad \text{and} \quad V^c(\lambda^*) \geq W^b[x^c(\lambda^*); \lambda^*]. \tag{3.35}
\]

The first inequality states that conditional on living in \( x^b(\lambda^*) \), the agent with ability \( \lambda^* \) has no incentive to switch to own a car. Since both car-owners and bus-users pay the same rent at \( x^b(\lambda^*) \), the inequality implies

\[
w\lambda^* - t\left[ w\lambda, x^b(\lambda^*) \right] \geq w\lambda - \tau\left[ w\lambda, x^b(\lambda^*) \right] - p_c w,
\]

or

\[
0 \geq t\left[ w\lambda, x^b(\lambda^*) \right] - \tau\left[ w\lambda, x^b(\lambda^*) \right] - p_c w. \tag{3.36}
\]

The second inequality in (3.35) can be interpreted similarly and implies

\[
t\left[ w\lambda, x^c(\lambda^*) \right] - \tau\left[ w\lambda, x^c(\lambda^*) \right] - p_c w \geq 0. \tag{3.37}
\]
Define a function $g : \mathbb{R}_+ \to \mathbb{R}$ by

$$g(x) = t(w\lambda, x) - \tau(w\lambda, x) - p_c w.$$  

Following the assumptions on $t(w\lambda, x)$ and $\tau(w\lambda, x)$, the function $g(\cdot)$ is continuous, differentiable with the first-order derivative given by

$$g'(x) = t_2(w\lambda, x) - \tau_2(w\lambda, x) > 0,$$

for all $x \geq 0$. This, together with (3.36) and (3.37) implies $x_c(\lambda) \geq x_b(\lambda)$. For reasons stated in the text, $x_c(\lambda) > x_b(\lambda)$ cannot hold in equilibrium. Hence, $x_c(\lambda) = x_b(\lambda)$. ■

**Numerical Algorithm**

The following is the algorithm for computing the model’s equilibrium:

1. Guess on a value of $\bar{\lambda}$. Compute $q(\lambda_{\text{min}})$ using (3.30) and (3.32).

2. Solve the bus-user’s initial value problem for $x_b(\lambda)$ and $q_b(\lambda)$.

3. Compute $\bar{x}$ using (3.16).

4. Update guess on $\bar{\lambda}$ and iterate until $|\bar{x} - x_b(\bar{\lambda})|$ is less than the desired tolerance.

5. To check the solution, solve the car owner’s initial value problem with initial conditions:

$$x_c(\bar{\lambda}) = \bar{x}, \quad \text{and} \quad q_c(\bar{\lambda}) = q_b(\bar{\lambda}),$$

for $x_c(\lambda)$ and $q_c(\lambda)$. Verify that $q_c(\lambda_{\text{max}})$ equals $q_A$. 

Chapter 3

3.8 Figures and Tables

Figure 3.1: Population Density Distributions for Atlanta in 1910, 1950, and 1970.

Figure 3.2: Population Density Distributions for Chicago in 1900 and 1950.
Figure 3.3: Number of Registered Automobiles per Capita, 1900-1993.

Figure 3.4: Percentage of Families who Own a Car by Number of Cars Owned, 1950-1995.
Figure 3.5: Car Ownership by Income Quintile, 1952-1965.

Figure 3.6: Time Cost of New Automobiles (in Logarithm), 1906-2000.
Figure 3.7: Car Ownership in 1962 by Distance from Center of Central City.

Figure 3.8: Car Ownership in 12 Largest SMSA’s by Place of Residence, 1961 and 1971.
Figure 3.9: Highway and Street Stock per Capita, 1925-94.

Figure 3.10: Capital Stock per Capita of Equipment and Structures for Intercity and Local Passenger Transit, 1947-1970.
Figure 3.11: Public Transit Ridership per Capita, 1902-1970.
Table 3.1
Coefficients for the Negative Exponential Function, \(d = ae^{-bx}\), for Chicago, 1900-1950.

<table>
<thead>
<tr>
<th>Year</th>
<th>a</th>
<th>b</th>
<th>% Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>1900</td>
<td>87,400</td>
<td>0.386</td>
<td></td>
</tr>
<tr>
<td>1910</td>
<td>92,900</td>
<td>0.344</td>
<td>-10.9</td>
</tr>
<tr>
<td>1920</td>
<td>90,400</td>
<td>0.297</td>
<td>-13.7</td>
</tr>
<tr>
<td>1930</td>
<td>84,700</td>
<td>0.256</td>
<td>-13.8</td>
</tr>
<tr>
<td>1940</td>
<td>78,850</td>
<td>0.233</td>
<td>-9.0</td>
</tr>
<tr>
<td>1950</td>
<td>69,700</td>
<td>0.201</td>
<td>-13.7</td>
</tr>
</tbody>
</table>


Table 3.2
Coefficients for the Negative Exponential Function, \(d = ae^{-bx}\), for Atlanta, 1900-1970.

<table>
<thead>
<tr>
<th>Year</th>
<th>a</th>
<th>b</th>
<th>% Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>1900</td>
<td>14,000</td>
<td>0.79</td>
<td></td>
</tr>
<tr>
<td>1910</td>
<td>29,000</td>
<td>0.94</td>
<td>19.0</td>
</tr>
<tr>
<td>1920</td>
<td>35,000</td>
<td>0.89</td>
<td>-5.3</td>
</tr>
<tr>
<td>1930</td>
<td>33,000</td>
<td>0.73</td>
<td>-18.0</td>
</tr>
<tr>
<td>1940</td>
<td>32,000</td>
<td>0.64</td>
<td>-12.3</td>
</tr>
<tr>
<td>1950</td>
<td>25,000</td>
<td>0.48</td>
<td>-25.0</td>
</tr>
<tr>
<td>1960</td>
<td>10,000</td>
<td>0.25</td>
<td>-47.9</td>
</tr>
<tr>
<td>1970</td>
<td>8,000</td>
<td>0.18</td>
<td>-28.0</td>
</tr>
</tbody>
</table>

Source: Edmonston (1975) p. 50.
Table 3.3
Mean Gradients for Forty-One U.S. Cities that were Metropolitan Districts in 1900.

<table>
<thead>
<tr>
<th>Year</th>
<th>b</th>
<th>% Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>1900</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>1910</td>
<td>0.96</td>
<td>-4.0</td>
</tr>
<tr>
<td>1920</td>
<td>0.86</td>
<td>-10.4</td>
</tr>
<tr>
<td>1930</td>
<td>0.62</td>
<td>-27.9</td>
</tr>
<tr>
<td>1940</td>
<td>0.57</td>
<td>-8.1</td>
</tr>
<tr>
<td>1950</td>
<td>0.45</td>
<td>-21.1</td>
</tr>
<tr>
<td>1960</td>
<td>0.34</td>
<td>-24.4</td>
</tr>
<tr>
<td>1970</td>
<td>0.28</td>
<td>-17.6</td>
</tr>
</tbody>
</table>


Table 3.4
Number of Registered Vehicles per Person Aged 20-64, 1900-1970.

<table>
<thead>
<tr>
<th>Year</th>
<th>Vehicle Per 100 Persons</th>
</tr>
</thead>
<tbody>
<tr>
<td>1900</td>
<td>0.02</td>
</tr>
<tr>
<td>1910</td>
<td>0.93</td>
</tr>
<tr>
<td>1920</td>
<td>14.12</td>
</tr>
<tr>
<td>1930</td>
<td>33.57</td>
</tr>
<tr>
<td>1940</td>
<td>35.39</td>
</tr>
<tr>
<td>1950</td>
<td>46.02</td>
</tr>
<tr>
<td>1960</td>
<td>65.66</td>
</tr>
<tr>
<td>1970</td>
<td>84.11</td>
</tr>
</tbody>
</table>

Table 3.5
Net Capital Expenditures of

<table>
<thead>
<tr>
<th>Year</th>
<th>Net Expenditures (millions in 1929 dollars)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1890</td>
<td>74.0</td>
</tr>
<tr>
<td>1895</td>
<td>176.2</td>
</tr>
<tr>
<td>1900</td>
<td>170.9</td>
</tr>
<tr>
<td>1905</td>
<td>229.8</td>
</tr>
<tr>
<td>1910</td>
<td>66.1</td>
</tr>
<tr>
<td>1915</td>
<td>15.2</td>
</tr>
<tr>
<td>1920</td>
<td>-128.5</td>
</tr>
<tr>
<td>1925</td>
<td>-105.4</td>
</tr>
<tr>
<td>1930</td>
<td>-85.3</td>
</tr>
<tr>
<td>1935</td>
<td>-60.7</td>
</tr>
<tr>
<td>1940</td>
<td>-10.4</td>
</tr>
<tr>
<td>1945</td>
<td>-58.4</td>
</tr>
<tr>
<td>1950</td>
<td>-53.5</td>
</tr>
</tbody>
</table>

Source: Ulmer (1960).
Table 3.6
Baseline Parameterization

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Income Distribution ($/year)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>7,134</td>
<td>9,214</td>
<td>10,005</td>
<td>15,226</td>
<td>19,469</td>
<td>24,835</td>
<td>28,246</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>3,716</td>
<td>4,805</td>
<td>5,216</td>
<td>7,938</td>
<td>10,151</td>
<td>12,949</td>
<td>14,726</td>
</tr>
<tr>
<td>Minimum Income Level</td>
<td>1,971</td>
<td>2,546</td>
<td>2,764</td>
<td>4,206</td>
<td>5,378</td>
<td>6,860</td>
<td>7,802</td>
</tr>
<tr>
<td>Maximum Income Level</td>
<td>19,907</td>
<td>25,715</td>
<td>27,918</td>
<td>42,486</td>
<td>54,330</td>
<td>69,303</td>
<td>78,814</td>
</tr>
<tr>
<td>Price of Car, $p_{c,t}$ ($)</td>
<td>32,880</td>
<td>14,934</td>
<td>9,412</td>
<td>9,664</td>
<td>14,253</td>
<td>15,399</td>
<td>12,271</td>
</tr>
<tr>
<td>Agricultural Land Rent, $q_{A,t}$ ($/Lot*)</td>
<td>7.29</td>
<td>4.86</td>
<td>3.21</td>
<td>5.23</td>
<td>13.89</td>
<td>11.23</td>
<td>19.55</td>
</tr>
<tr>
<td><strong>Transportation Costs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Bus</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fixed Cost, $\gamma_{b,t}$ ($/day)</td>
<td>0.05</td>
<td>0.11</td>
<td>0.25</td>
<td>0.58</td>
<td>1.33</td>
<td>3.05</td>
<td>6.97</td>
</tr>
<tr>
<td>Time Cost, $\tau_{b,t}$ (hours/mile)</td>
<td>0.46</td>
<td>0.77</td>
<td>1.30</td>
<td>1.53</td>
<td>1.86</td>
<td>2.91</td>
<td>4.08</td>
</tr>
<tr>
<td>Average Distance (miles)</td>
<td>2.32</td>
<td>1.01</td>
<td>0.45</td>
<td>0.31</td>
<td>0.26</td>
<td>0.15</td>
<td>0.08</td>
</tr>
<tr>
<td>Hourly Wage at Average Distance</td>
<td>5.61</td>
<td>5.61</td>
<td>4.56</td>
<td>5.55</td>
<td>7.15</td>
<td>7.59</td>
<td>6.70</td>
</tr>
<tr>
<td>Average Variable Cost ($/day)</td>
<td>5.97</td>
<td>4.35</td>
<td>2.64</td>
<td>2.02</td>
<td>3.47</td>
<td>3.25</td>
<td>2.28</td>
</tr>
<tr>
<td>Average Total Cost ($/day)</td>
<td>6.02</td>
<td>4.46</td>
<td>2.90</td>
<td>3.20</td>
<td>4.81</td>
<td>6.30</td>
<td>9.25</td>
</tr>
<tr>
<td><strong>Car</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fixed Cost, $\gamma_{c,t}$ ($/day)</td>
<td>5.88</td>
<td>2.44</td>
<td>4.06</td>
<td>4.87</td>
<td>6.69</td>
<td>7.19</td>
<td>8.11</td>
</tr>
<tr>
<td>Time Cost, $\tau_{c,t}$ (hours/mile)</td>
<td>0.46</td>
<td>0.25</td>
<td>0.20</td>
<td>0.16</td>
<td>0.15</td>
<td>0.14</td>
<td>0.11</td>
</tr>
<tr>
<td>Average Distance (miles)</td>
<td>—</td>
<td>5.02</td>
<td>5.46</td>
<td>6.86</td>
<td>6.19</td>
<td>6.95</td>
<td>8.13</td>
</tr>
<tr>
<td>Hourly Wage at Average Distance</td>
<td>—</td>
<td>10.84</td>
<td>10.43</td>
<td>14.64</td>
<td>18.48</td>
<td>22.53</td>
<td>23.79</td>
</tr>
<tr>
<td>Average Variable Cost ($/day)</td>
<td>—</td>
<td>13.44</td>
<td>11.51</td>
<td>15.77</td>
<td>17.66</td>
<td>22.52</td>
<td>21.76</td>
</tr>
<tr>
<td>Average Total Cost ($/day)</td>
<td>—</td>
<td>15.88</td>
<td>15.57</td>
<td>20.64</td>
<td>24.35</td>
<td>29.71</td>
<td>29.88</td>
</tr>
</tbody>
</table>

* One lot is 12,910 square feet.
### Table 3.7  
Baseline Results

<table>
<thead>
<tr>
<th>Year</th>
<th>Car Ownership (%)</th>
<th>Population density gradient</th>
<th>% change in gradient</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Data</td>
<td>Model</td>
<td>Data</td>
</tr>
<tr>
<td>1910</td>
<td>0.93</td>
<td>0.83</td>
<td>—</td>
</tr>
<tr>
<td>1920</td>
<td>14.12</td>
<td>0.79</td>
<td>-4.8</td>
</tr>
<tr>
<td>1930</td>
<td>33.57</td>
<td>0.66</td>
<td>-16.5</td>
</tr>
<tr>
<td>1940</td>
<td>35.39</td>
<td>0.61</td>
<td>-7.6</td>
</tr>
<tr>
<td>1950</td>
<td>46.02</td>
<td>0.39</td>
<td>-36.1</td>
</tr>
<tr>
<td>1960</td>
<td>65.66</td>
<td>0.31</td>
<td>-20.5</td>
</tr>
<tr>
<td>1970</td>
<td>84.11</td>
<td>0.23</td>
<td>-25.8</td>
</tr>
</tbody>
</table>

Model values are in parentheses.

<table>
<thead>
<tr>
<th>Year</th>
<th>End of Bus-users (miles)</th>
<th>Boundary of City (miles)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1910</td>
<td>6.25</td>
<td>6.25</td>
</tr>
<tr>
<td>1920</td>
<td>2.91</td>
<td>7.98</td>
</tr>
<tr>
<td>1930</td>
<td>1.47</td>
<td>11.69</td>
</tr>
<tr>
<td>1940</td>
<td>1.08</td>
<td>15.34</td>
</tr>
<tr>
<td>1950</td>
<td>0.94</td>
<td>12.70</td>
</tr>
<tr>
<td>1960</td>
<td>0.56</td>
<td>15.83</td>
</tr>
<tr>
<td>1970</td>
<td>0.31</td>
<td>17.90</td>
</tr>
</tbody>
</table>

* The values of R-squared are in parenthesis.
Table 3.8
Counterfactual Experiments

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Car-Owners</td>
<td>0.93</td>
<td>14.12</td>
<td>33.57</td>
<td>35.39</td>
<td>46.02</td>
<td>65.66</td>
<td>84.11</td>
</tr>
<tr>
<td>Density Gradient</td>
<td>0.83</td>
<td>0.79</td>
<td>0.66</td>
<td>0.61</td>
<td>0.39</td>
<td>0.31</td>
<td>0.23</td>
</tr>
<tr>
<td><strong>2. Baseline Model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Car-Owners</td>
<td>0</td>
<td>5.36</td>
<td>19.09</td>
<td>25.52</td>
<td>35.22</td>
<td>51.27</td>
<td>72.47</td>
</tr>
<tr>
<td>Density Gradient</td>
<td>0.86</td>
<td>0.83</td>
<td>0.63</td>
<td>0.46</td>
<td>0.51</td>
<td>0.46</td>
<td>0.37</td>
</tr>
<tr>
<td><strong>3. Prices and Wages Remain at 1910 Values</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a) With rising cost of public transportation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Car-Owners</td>
<td>0</td>
<td>0</td>
<td>0.20</td>
<td>0.27</td>
<td>0</td>
<td>0.34</td>
<td>0.52</td>
</tr>
<tr>
<td>Density Gradient</td>
<td>0.86</td>
<td>1.42</td>
<td>2.05</td>
<td>2.38</td>
<td>3.58</td>
<td>4.62</td>
<td>6.79</td>
</tr>
<tr>
<td>b) Without rising cost of public transportation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Car-Owners</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Density Gradient</td>
<td>0.86</td>
<td>0.85</td>
<td>0.84</td>
<td>0.85</td>
<td>0.88</td>
<td>0.87</td>
<td>0.90</td>
</tr>
<tr>
<td><strong>4. Price of Car Remains at 1910 Value</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a) With rising cost of public transportation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Car-Owners</td>
<td>0</td>
<td>0.19</td>
<td>2.43</td>
<td>9.93</td>
<td>15.88</td>
<td>29.65</td>
<td>43.09</td>
</tr>
<tr>
<td>Density Gradient</td>
<td>0.86</td>
<td>1.29</td>
<td>1.13</td>
<td>0.68</td>
<td>0.67</td>
<td>0.55</td>
<td>0.47</td>
</tr>
<tr>
<td>b) Without rising cost of public transportation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Car-Owners</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2.22</td>
<td>3.57</td>
<td>7.23</td>
<td>11.20</td>
</tr>
<tr>
<td>Density Gradient</td>
<td>0.86</td>
<td>0.85</td>
<td>0.84</td>
<td>0.85</td>
<td>0.88</td>
<td>0.87</td>
<td>0.90</td>
</tr>
<tr>
<td><strong>5. Wage Distribution Remains as in 1910</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a) With rising cost of public transportation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Car-Owners</td>
<td>0</td>
<td>2.65</td>
<td>11.20</td>
<td>12.92</td>
<td>6.21</td>
<td>8.51</td>
<td>17.30</td>
</tr>
<tr>
<td>Density Gradient</td>
<td>0.86</td>
<td>0.97</td>
<td>0.74</td>
<td>0.70</td>
<td>1.27</td>
<td>1.23</td>
<td>1.18</td>
</tr>
<tr>
<td>b) Without rising cost of public transportation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Car-Owners</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Density Gradient</td>
<td>0.86</td>
<td>0.85</td>
<td>0.65</td>
<td>0.61</td>
<td>0.88</td>
<td>0.87</td>
<td>0.90</td>
</tr>
<tr>
<td><strong>6. Stock of Highways and Roads per Capita Remains as in 1910</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Car-Owners</td>
<td>0</td>
<td>0.65</td>
<td>8.63</td>
<td>17.80</td>
<td>18.86</td>
<td>33.34</td>
<td>53.62</td>
</tr>
<tr>
<td>Density Gradient</td>
<td>0.86</td>
<td>1.26</td>
<td>1.26</td>
<td>1.18</td>
<td>1.27</td>
<td>1.23</td>
<td>1.22</td>
</tr>
<tr>
<td><strong>7. Time and Fixed Cost of Public Transit Remains at 1910 Value</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Car-Owners</td>
<td>0</td>
<td>1.10</td>
<td>5.52</td>
<td>14.44</td>
<td>12.06</td>
<td>16.92</td>
<td>26.19</td>
</tr>
<tr>
<td>Density Gradient</td>
<td>0.86</td>
<td>0.73</td>
<td>0.56</td>
<td>0.42</td>
<td>0.46</td>
<td>0.39</td>
<td>0.32</td>
</tr>
<tr>
<td><strong>8. Stock of Highways and Roads per Capita, and Cost of Using Public Transportation at 1910 Values</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Car-Owners</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Density Gradient</td>
<td>0.86</td>
<td>0.85</td>
<td>0.84</td>
<td>0.84</td>
<td>0.85</td>
<td>0.84</td>
<td>0.85</td>
</tr>
</tbody>
</table>


