THREE ESSAYS ON PUBLIC POLICY, HUMAN CAPITAL, AND ECONOMIC GROWTH: THEORY AND EVIDENCE

Gregory A. W. Gilpin

Submitted to the faculty of the University Graduate School
in partial fulfillment of the requirements
for the degree
Doctor of Philosophy
in the Department of Economics,
Indiana University
August 2009
Accepted by the Graduate Faculty, Indiana University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Doctoral Committee

______________________________________
Michael Kaganovich, Chair

______________________________________
William E. Becker

______________________________________
Gerhard Glomm

______________________________________
Rusty Tchernis

Date of Oral Examination: August 11, 2009
© (2009)
Gregory A. W. Gilpin
ALL RIGHTS RESERVED
To my beautiful wife, Renée, and our four loving children,
Julia, Tyler, Ethan, and Jacob

Their love and support
made this happen.
I am indebted to Michael Kaganovich and Gerhard Glomm for their guidance and support over this process. I would like to thank the other members of my committee, William E. Becker and Rusty Tchernis, and the faculty of the Economics department at Indiana University for their constant support. I acknowledge financial support from the Economics Department at Indiana University, the College of Arts and Sciences at Indiana University, the Spencer Foundation, and the Bill and Melinda Gates Foundation.
This dissertation presents three essays investigating the effect of public policies on human capital accumulation and economic growth. I first investigate the optimal quantity and quality of teachers in the framework of dynamic general equilibrium OLG model. The government hires teachers while constrained by the teachers’ collective bargaining agreement. In the process of endogenous growth, the optimal trade-off between the quantity and quality of teachers moves in the direction of the former. The number of teachers hired will grow over time while their relative human capital attainment will fall. This evolution of human capital accumulation is accompanied by increasing inequality, within the group of college educated workers in particular.

In the second essay, I empirically investigate the self-selection of first year teachers into teaching positions at union and non-union schools, the main difference being the entry and dismissal barriers of teachers’ unions. In a Bayesian framework with Markov Chain Monte Carlo methods, I estimate how much of the propensity to remain in teaching can be attributed to the effect of teachers’ unions rather than the characteristics of individuals who tend to work in unionized schools. With the use of the counterfactual, I find that while teachers’ unions do lower the attrition rate, most of the difference in
attrition rates should be attributed to the self-selection into the teaching occupation and not to the barriers to dismissal unions provide.

The third essay analyzes the implications of pay-as-you-go (PAYG) social security system for human capital investment, economic growth and income distribution when individuals’ longevities are dependent on their human capital attainment. While PAYG pensions encourage human capital investment at a cost of lower growth in the short-run, it results in increased life expectancy of the subsequent generations, providing additional incentive to save for retirement. Thus, in the long-run, implementing PAYG social security causes the stock of human capital to grow; savings will increase due to the dynamic externality, thereby so will the aggregate output.
Table of Contents

Chapter 1

The Quantity and Quality of Teachers: A Dynamic Trade-off

Abstract ..........................................................................................................................1

1. Introduction ..............................................................................................................2

2. The Model ...............................................................................................................7
   2.1 Production .........................................................................................................8
   2.2 Households ......................................................................................................8
   2.3 Human Capital Formation .............................................................................9
   2.4 Quality of Basic Education ...........................................................................12
   2.5 Government ...................................................................................................14

3. General Equilibrium and Optimal Policy ...............................................................18

4. Analysis of the Model ...........................................................................................21

5. Conclusion .............................................................................................................33

References ..................................................................................................................34

Appendices ..................................................................................................................37

   Selected Proofs .......................................................................................................37

   Glossary of Mathematical Terms .........................................................................43
# Chapter 2

*Self-selection, Unions, and Teacher Attrition*

Abstract .................................................................................................................. 45

1. Introduction ....................................................................................................... 46

2. The Determinants of First Year Teachers’ Mobility Decision ......................... 49

3. The Data ............................................................................................................ 54

4. The Model ......................................................................................................... 61

   4.1 The Likelihood .................................................................................... 65

   4.2 Bayesian Estimation ............................................................................ 67

   4.3 Treatment Effects ................................................................................ 68

5. Results ............................................................................................................... 70

6. Conclusion ........................................................................................................ 76

References .............................................................................................................. 80

Electronic References ............................................................................................ 83

Appendices ............................................................................................................. 84

   Data Construction .......................................................................................... 84

   The Posterior Simulator and Data Generated Experiment .......................... 90

   Associated Tables .......................................................................................... 96
Chapter 3

Unequal Longevity and Social Security: The Implications for Human Capital, Growth and Distribution

Abstract ........................................................................................................................................98

1. Introduction................................................................................................................................99

2. The Model....................................................................................................................................102
   2.1 Human Capital Formation........................................................................................................103
   2.2 Production.................................................................................................................................103
   2.3 Longevity................................................................................................................................103
   2.4 Government..............................................................................................................................105
   2.5 Preferences and Budget Constraints .........................................................................................106
   2.6 Individual’s Maximization Problem.........................................................................................106

3. Definition of Dynamic Competitive Equilibrium.......................................................................107

4. Solving the Model......................................................................................................................108

5. Analysis of the Model...............................................................................................................112

6. Conclusion..................................................................................................................................117

References.....................................................................................................................................119

Appendices....................................................................................................................................123

Numerical Result 5.1. ....................................................................................................................123
Numerical Result 5.2. ....................................................................................................................126
List of Tables and Figures

Chapter 1

Table 1 Historical Data on Public Elementary and Secondary Schools from 1955:2005 .................................................................2

Figure 1 Basic Education Optimization Problem .................................................17
Figure 2 Optimal Education Policy at $t$ and $t+1$ ..................................................29

Chapter 2

Table 1 Mobility Rates Among Teachers With 1 Year Experience ......................55
Table 2 Schools’ Barriers to Dismissal of Teachers.............................................56
Table 3 Most Important Reasons Why Teachers Say They Leave .......................57
Table 4 Occupational Status of Former Teachers With 1 Year Experience ...........58
Table 5 Average Earnings of Former First Year Teachers....................................59
Table 6 Average Earnings of First Year Teachers Who Remain Teaching ..............60
Table 7 Certification Type of First Year Teachers ..............................................60
Table 8 Expected Return Date of Former Teachers.............................................61
Table 9 Marginal Effects from Treatment Equation and Correlation Estimates .71
Table 10 Marginal Effects of Outcome Equations ................................................72
Table 11 ATE from Unionization on Teacher Attrition .......................................75
Table 12 The Effect of Dismissal Barriers on Teacher Attrition ............................76
Table 13 Union and Non-Union Schools Teacher Attrition Over Time...............78
Table A1 Variable Labels, Definitions, and Date Sources ...........................................87
Table A2 Ancillary Data Summary .................................................................................88
Table A3 Summary Statistics ......................................................................................89
Table B1 Data Generated Simulation Results ..........................................................95
Table C1 NAICS 2000 Occupation Codes ..................................................................96
Table C2 Tenure and Collective Bargaining State Summary ......................................97

Figure 1 An Individual’s Propensity to Exit ............................................................64
Figure 2 Union Effect on Teacher Attrition ............................................................70
Figure 3 Individuals’ Propensities to Exit Teaching ..................................................74
Figure 4 Union Effect on Exit Decision .................................................................75

Chapter 3

Table A1 Numerical Result 5.1. .............................................................................123
Table B1 Numerical Result 5.2. .............................................................................126
Chapter 1

The Quantity and Quality of Teachers: A Dynamic Trade-off

Abstract

We study the dynamics of the quantity and quality of teachers in the framework of dynamic general equilibrium OLG model. The quantity and quality are jointly set by a government agency wishing to maximize the quality of basic education per student while bound by teachers’ collective bargaining agreement which equalizes teacher pay. Our model features two stages of education: basic and advanced (college), the latter being required of teachers. The cost of hiring teachers is influenced by the outside opportunities that college graduates have in the production sector. We show that this factor strengthens in the process of endogenous growth and moreover that it pushes the optimal trade-off between quantity and quality of teachers in the direction of the former. Namely, the number of teachers hired will grow over time while their relative (but not the absolute) human capital attainment will fall. This evolution of human capital accumulation is accompanied by increasing inequality, within the group of college educated workers in particular. Further, we consider the comparative dynamics effect of exogenous skill biased technological change represented by a positive shock to productivity of the skilled workers, hence to the college premium. We show that this will exacerbate the negative trends in the quality of basic education in relation to GDP growth. Countering this trend would therefore require an increase in the share of GDP spent on basic education, assuming that the institutional setup of the school system remains unchanged.

JEL classification: H52, I2, O4
1. Introduction

Increasing focus on “individual based instruction” continues to be one of the main education policy priorities in the United States as a means to raising education quality. This is evidenced by the dynamics of student-teacher ratio which has fallen from 25.8 in 1960 to 15.7 in 2005 (Digest for Education Statistics 2007, table 61). Research, however, has shown that students' test scores have not risen despite increased individualized instruction. This discrepancy has compelled policy makers and researchers to question the factors affecting students' test scores and the role of the quality of teachers vs. their quantity (see Hanushek et al (2005)). This paper develops a theoretical framework for analyzing this quantity-quality trade-off and offers an explanation to the observed trend biased in favor of quantity.

Some of the changes in education statistics between 1955 and 2005 are displayed in the table below. It shows that the decline in the student-teacher ratio was accompanied by declining relative teacher salaries while the overall K-12 public education expenditures have been essentially flat as a share of GDP since 1970.

<table>
<thead>
<tr>
<th>Year</th>
<th>Enrollment</th>
<th>Teachers</th>
<th>Pupil / teacher ratio</th>
<th>Expenditures to GDP</th>
<th>Relative Teacher Salary</th>
</tr>
</thead>
<tbody>
<tr>
<td>1955</td>
<td>30,680</td>
<td>1,141</td>
<td>26.9</td>
<td>3.3</td>
<td></td>
</tr>
<tr>
<td>1960</td>
<td>36,281</td>
<td>1,408</td>
<td>25.8</td>
<td>3.6</td>
<td>43</td>
</tr>
<tr>
<td>1965</td>
<td>42,173</td>
<td>1,710</td>
<td>24.7</td>
<td>3.9</td>
<td>44</td>
</tr>
<tr>
<td>1970</td>
<td>45,894</td>
<td>2,059</td>
<td>22.3</td>
<td>4.6</td>
<td></td>
</tr>
<tr>
<td>1975</td>
<td>44,819</td>
<td>2,198</td>
<td>20.4</td>
<td>4.6</td>
<td></td>
</tr>
<tr>
<td>1980</td>
<td>40,877</td>
<td>2,184</td>
<td>18.7</td>
<td>4.0</td>
<td>41</td>
</tr>
<tr>
<td>1985</td>
<td>39,422</td>
<td>2,206</td>
<td>17.9</td>
<td>3.8</td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td>41,217</td>
<td>2,398</td>
<td>17.2</td>
<td>4.3</td>
<td>35</td>
</tr>
<tr>
<td>1995</td>
<td>44,840</td>
<td>2,598</td>
<td>17.3</td>
<td>4.3</td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>47,204</td>
<td>2,941</td>
<td>16.0</td>
<td>4.5</td>
<td>36.5</td>
</tr>
<tr>
<td>2005</td>
<td>49,113</td>
<td>3,137</td>
<td>15.7</td>
<td>4.6</td>
<td></td>
</tr>
</tbody>
</table>

Source: 1: Digest for Education Statistics 2007, Table 61  
2: Digest for Education Statistics 2007, Table 25  
Notes: a: In thousands  
b: In Percent  
c: 1959 data  
d: College educated females, age 20-29, earning less than average female teacher, age 20-29

Another significant trend observed over about the same period is the decline in the aptitude of teachers relative to other educated workers. Hoxby and Andrews (2004) estimate that in 1963 41% of all teachers were of the “middle” aptitude relative to their educated peers, with 17% above and 42% below the average; by comparison, in 2000,
28% of all teachers were of the “middle” aptitude with 5% above and 67% below average. Corcoran et al. (2002) provide similar results. Notably, student test scores have remained roughly constant despite substantial growth in per student public education outlays. Much attention in the literature has focused on explaining how these different inputs in K-12 public school system have affected students' test scores.

Many of the conflicting conclusions in the literature concerning the factors affecting student performance stem from two general empirical strategies employed in order to estimate the returns to quality and quantity of teachers. The first strategy attempts to estimate the effect of teacher characteristics on student achievement while partially controlling the class size (see Aaronson (2007), Clotfelter (2007), Rivkin et al. (2005), Goldhaber and Anthony (2007)). Class size is naturally constrained by geographic and time specification of the observations (schools in the same state are under one mandated student-teacher ratio). The second empirical strategy aims to estimate how class size affects student achievement while attempting to control for teacher quality. Several studies which follow this strategy use data from policy experiments producing random assignment of students to smaller and larger classes. Controlling for teacher quality this data yields unbiased estimates of the effects of class size on student achievement (see Angrist and Lavy (1999), Krueger and Whitmore (2001), Krueger (1999), Jepsen and Rivkin (2002)).

Using data from North Carolina, Clotfelter et al. (2007) conclude that teacher experience, test scores and regular licensure all have greater positive effects on student achievement, whether compared to the effects of changes in class size or to the socioeconomic characteristics of students. Aaronson et al. (2007) use the data on Chicago public high school students and teachers at the classroom level to estimate how teacher characteristics affect mathematics test scores. They find that replacing a teacher with another who is rated two standard deviations superior in quality can add 0.35 to 0.45 grade equivalents, or 30 to 40 percent of an average school year’s worth, to a student's math score performance. Goldhaber and Anthony (2007) use the same North Carolina data to examine the effects of the National Board Certification of teachers and find mixed evidence on whether improved observable teacher credentials have positive impact on
student achievement. These results are similar to Rivkin et al. (2005) who use the UTD Texas Schools Project. On the other hand, Angrist and Lavy (1999) use Israel’s class size maximum to estimate class size effects on student achievement. They find that reducing class size causes significant and substantial increase in test scores for fourth and fifth graders, although not for third graders. Krueger (1999) analyzes data from Tennessee Project STAR to estimate the effects of class size reductions on student performance on standardized tests. His results indicate that students’ scores increase by four percentage points in the first year they attend smaller classes while in subsequent years the test scores grow by about one percentage point per year. Hanushek (1999) rebuts Krueger’s findings citing important design and implementation issues in the STAR project suggesting an upward bias in the returns to class size reduction. Krueger and Whitmore (2001) follow up on students who participated in the Tennessee STAR experiment and find that they had, on average, ACT scores of .13 standard deviations higher.

Another approach uses longitudinal data on declining class size. Card and Krueger (1992) find that a decrease in the pupil-teacher ratio from 30 to 25 is associated with a 0.4 percentage point increase in the rate of return to education. Hoxby (2000), however, estimates that there is no effect from decreased class size on student achievement. These opposing estimates are addressed by Jepson and Rivkin (2002). They argue that using mandated class size reduction programs as natural experiments for estimating the class size effect is problematic when these changes involve a trade-off between the quantity and quality of teachers, and that the same problem arises when time series data is used without the account for this endogenous trade-off. Specifically, their results indicate that California’s class size reduction program came at a cost of hiring lower quality teachers to staff additional classrooms which offset the benefits of smaller classes. Similarly, Hoxby (1996) finds that in unionized public school higher measured inputs may produce no gains to student achievement: the unions engage in rent-seeking which leads to lower productivity, via teacher quality or effort, enough to offset any gains from the additional resources, smaller student-teacher ratios in particular.

Thus, despite a significant attention in the literature, the questions about the determinants of education quality remain open. This underscores the need for a broader
theoretical framework, which would capture the dynamic interaction between inputs in education as it is influenced by labor market in the production economy. We note in this regard a branch of recent literature which has studied how outside job market opportunities have affected the quality of teachers. Flyer and Rosen (1997) report that the three-fold increase in direct costs of education per student is attributable to the growing market opportunities for women. Hanushek and Rivkin (1997) document the decline in the earnings of women teachers relative to women in other occupations and suggest that the expansion of alternative opportunities reduced teacher quality. Hanushek and Rivkin (2003) estimate that in 1955, 50% of all educated male workers earned less than male teachers, compared to 36% in 2000. Likewise, in 1955, 48% of all educated female workers earned less than female teachers compared to 29% in 2000. Similar analyses concerning the effect of the outside work opportunities on teacher quality are offered by Goldhaber and Liu (2003), Stoddard (2003), and Bacolad (2006). Lakdawalla (2006) demonstrates that a rising skill premium of educated workers due to faster technical change coupled with low productivity growth of skilled teachers, has lead to lower teacher quality. The mechanism he highlights is the substitution of unskilled teachers for increasingly expensive skilled teachers.

In this paper we present a theoretical model which incorporates some of the factors of education quality discussed above, in a dynamic general equilibrium framework where government education policy decisions affect and are affected by individual education and employment decisions, whereas the dynamics of human capital accumulation and labor productivity has a feedback effect on both. In our model, the government agency wishes to maximize the quality of basic education per student and faces a trade-off between the quality and quantity of teachers to be hired. Furthermore, we assume that the agency is bound by teachers‘ collective bargaining agreement which equalizes teacher pay. It is, indeed, well documented that teachers‘ unions significantly contribute to the wage compression phenomenon. Unions provide tenure to teachers and tie their salary primarily to experience rather than performance. Administrators wishing to hire higher quality teachers are forced by the unions to then provide matching raises to
teachers across the board.\textsuperscript{1} Furthermore, the wage uniformity in public schools imposes similar wage rigidity on the private school teacher market (Lakdawalla, 2006).

In our model, a government education agency has two policy variables: teacher salary, which is uniform according to the collective bargaining regime, and the number of teachers to be hired. The model features two stages of education: basic and advanced (college), the latter being required of teachers. College graduates can also take jobs in the skilled labor force of the production sector and get paid a competitive wage according to their human capital attainment. This opportunity cost implies that the level of teacher salary set by the government agency will determine the top quality (human capital level) of a teacher who can be hired at this salary. All college graduates whose human capital is below this level will be motivated to take a teaching job at the same salary. Therefore, given the top quality cut-off determined by the government-set teacher salary, the number of teachers the government decides to hire will determine the lowest human capital cut-off among teachers. Thus the total cost of hiring teachers is affected in our model by the outside opportunities available to skilled individuals in the production sector. We show, moreover, that in the process of endogenous growth this effect strengthens and that it pushes the optimal trade-off between quantity and quality of teachers in the direction of the former. Namely, in the face of rising over time cost of highly able skilled workers, the government agency will find it optimal to opt for increasing the number of teachers hired while reducing the overall relative quality of the pool of teachers. (The absolute human capital attainment of teachers, however, will rise along with the overall human capital accumulation, while sliding toward the lower tail of the distribution of college educated population.) Furthermore, we show that this human capital dynamics is characterized by increasing inequality within the group of college educated workers as well as between it and the unskilled.

\textsuperscript{1} It should be noted that unionization is not the sole explanation for the compression of teacher salaries. It is also due in part to the difficulty of measuring teacher productivity, especially in terms of educational value added given unobservable student characteristics. But even if such characteristics were observable, there still exists the challenge of determining criteria for performance based pay for teachers. For example, low ability students exhibit relatively low average gains in learning throughout the year, therefore an approach based on marginal improvement of students’ performance would not fairly compensate teachers for working with lower ability children.
Thus this paper offers a theory explaining the trend in education policy in favor of lower student-teacher ratios (i.e., higher quantity of teachers) combined arguably with deteriorating teacher quality, despite growing per student schooling expenditures. We then build on these results to further analyze the impact of exogenous technological change biased toward skill, i.e., augmenting productivity of skilled workers and thereby the college premium. We show that such technological change will exacerbate the negative trends in the quality of basic education in relation to GDP growth. Specifically, the comparative dynamics effect (relative to the benchmark trajectory) will be detrimental to the aggregate quality of teachers as well as to the quality of basic education, due to the upward shocks to the cost of skilled labor.

The paper is organized as follows. Section 2 develops a dynamic general equilibrium model with unionized public schools. Section 3 defines a competitive equilibrium, Section 4 derives main results, and Section 5 concludes. Appendix 1 contains some of the more technical the proofs. Appendix 2 contains a glossary of notation.

2. The Model

We develop a general equilibrium growth model of an economy populated by overlapping generations of individuals whose life consists of three periods: childhood, young adulthood, and old age. We identify a generation with the period when its members are young adults, thus the individuals born in period $t-1$ form a generation $G_t$.

We assume that population size is constant in each generation $G_t$ and that it forms a continuum on the interval $[0,1]$. Let $\mu(.)$ be the induced Lebesgue measure on this set, so that $\mu([0,1])=1$ for all $t$.

Children make no decisions of their own and receive basic (or first stage) education which is provided publicly. Young adults are endowed with a unit of time and face an option of devoting a fixed fraction $n$ of it to acquiring advanced education (which we will also refer to as college or second stage education); the balance of time not
spent on education is inelastically devoted to work. Specifically, the individuals without college education will work for the full unit of time in the “unskilled” production workforce. Those with college education either work for the remaining fraction of time $1-n$ in the “skilled” production workforce or, if qualified by the government, can work as public school teachers. Individuals derive income from work. They spend part of it on consumption when young and invest the rest to use the returns to finance their consumption in retirement, the last period of life.

2.1. Production
The production sector of the economy consists of private perfectly competitive firms producing a homogeneous capital/consumption good by means of a constant returns technology which uses three factors of production - physical capital as well as unskilled and skilled human capital. The aggregate production function is given by

$$Y_t = D K_t^\alpha \left[ H_t^u + \theta_t H_t^{sy} \right]^{1-\alpha},$$  \hspace{1cm} (1)

where $\alpha \in [0,1]$, $D > 0$, while $K_t$, $H_t^u$, $H_t^{sy}$ stand, respectively, for aggregate supply of physical capital, unskilled human capital, and skilled human capital employed in the production sector in period $t$. The coefficient $\theta_t$ characterizes the net productivity augmentation of skilled human capital (adjusted for the shorter employment duration due to the time spent in college) which is imbedded in technology. The sequence of \{\theta_t\}_{t=0}^\infty characterizing the evolution of the skill premium in the process of technological change is assumed to be exogenously given.\footnote{See Appendix 2 for the glossary of notation.}

2.2. Households
All individuals $\omega$ of generation $G$, $t = 0,1,2,...$ have identical intertemporal preferences over consumption as young adults and retirees given by

$$\ln c_{t,t}^{\omega} + \beta \ln c_{t,t+1}^{\omega}$$  \hspace{1cm} (2)

subject to the life-time budget constraint.
where \( r_{t+1} \) is the market interest rate, \( I_t(\omega) \) is the individual’s wage income derived from human capital, while \( \tau_t \) is the uniform rate of labor income tax collected by the government. According to the production function (1) individuals working in the production economy receive the wage at competitive rates \( w_t \) and \( \theta_t w_t \), respectively, per unit of their unskilled or skilled human capital, whichever applies. Thus the income of individual \( \omega \) who receives only basic education and attains the level of unskilled human capital \( h^u_t(\omega) \) will be

\[
I^u_t(\omega) = w_t h^u_t(\omega)
\] (4)

The individual \( \omega \) who obtains college education, attains the level of skilled human capital \( h^s_t(\omega) \) and is employed in the production sector, will receive income

\[
I^s_t(\omega) = \theta_t w_t h^s_t(\omega)
\] (5)

College educated individuals who become teachers will receive income \( I^t_t \) to be specified later.

2.3. Human Capital Formation

The human capital received by each child \( \omega \) of generation \( G_t \) at the first (basic) stage of his education is produced in period \( t-1 \) by combining children’s random innate ability with public education according to

\[
h^u_t(\omega) = C a(\omega) E_{t-1}
\] (6)

where \( C \) is a positive constant, \( E_{t-1} \) is a uniform quality of public schooling received by each child in period \( t-1 \) while \( a(\omega) \) is the child’s innate ability. We assume that innate ability is distributed independently and identically in each generation (the time indexation is thus omitted); specifically, the distribution is uniform on the interval \([a, A]\). To simplify the exposition (but at no cost to the essence of the matter) we will later let \( a = 0 \).

We will now introduce the human capital production function for the advanced (college) stage of education. Consistent with Ben-Porath (1967) and Rosen (1976) we
assume that the gains from college education depend on one’s prior preparation, which in turn depends on innate ability. Moreover, we assume that college education has a pre-requisite human capital threshold $h^*$. Rather than an *ad hoc* admission requirement (we assume that all individuals are free to choose to go to college but base this decision purely on income considerations) we view this threshold as a set of benchmark skills, such as adequate language and mathematical proficiency whose deficit would preclude any benefit from learning at the advanced stage.$^3$ Specifically, we postulate that if an individual $\omega$ of generation $G_t$ chooses to go to college, he will become a “skilled” agent with the level of human capital given by

$$ h_t^s(\omega) = bh_t^u(\omega) + B[h_t^u(\omega) - h^*] $$

(7)

where $b \in (0,1)$ and $B > 0$ are given constants. Thus according to the expression (7) the gains from college education depend on the extent to which the individual’s pre-college human capital attainment $h_t^u(\omega)$ exceeds the threshold $h^*$.

The college education production function (7) also reflects a partial loss of pre-college human capital, according to the coefficient $b$, for the purposes of skilled human capital. While this loss is counteracted by the net productivity augmentation $\theta_t$ of skilled human capital according to the economy’s production function (1), we impose a condition

$$ b\theta_t < 1 $$

(8)

which indeed implies that individuals whose pre-college human capital $h_t^u(\omega)$ is below, at, or even slightly above the threshold $h^*$, will not gain from attending college and therefore will not choose to do so. It is likewise logical to assume that highly able individuals, particularly those with the highest ability level $A$, will benefit from attending

---

$^3$ See Su (2004) for a similar approach to college eligibility. One can envision that this knowledge threshold may evolve over time with changes in learning technology. For example, it now tends to incorporate computer literacy; while applicants are not tested on it for admission decisions, their progress in many college specialties will critically depend on it. For the purposes of our analysis $h^*$ is assumed fixed. Note the important distinction between this knowledge threshold, which determines a student’s true performance, and the concept of an *ad hoc* admission threshold addressed in the “educational standards” literature, such as Betts (1998), an education policy variable that serves as a sorting device and employability signal.
college. According to equations (6) and (7) this will be true in generation $G_t$ iff the following inequality holds:

$$(b + B)CAE_{t-1} \geq Bh^*$$

In Section 4 we will state specific parametric assumptions which in particular ensure that the above inequality does hold at all times.

According to the expressions (6) and (7) human capital of each type, and therefore the corresponding income, is an increasing function of the innate ability. Therefore if a certain individual decides to attend college then all agents with higher ability will also do so. Thus in each period $t$ there is an ability cut-off level $a^*_t$ such that an individual $\omega$ in generation $t$ will choose to attend college if and only if his ability $a(\omega)$ exceeds $a^*_t$. (Without loss of generality we’ll make a convention that individuals with ability level $a^*_t$ do choose to go to college.) We will later show that this college attendance ability cut-off level is given by the formula

$$a^*_t = \frac{1}{CE_{t-1}} \frac{\theta_t Bh^*}{\theta_t(b + B) - 1}$$

(9)

which has a straightforward meaning: an individual will choose to attend college if and only if his wage income derived from skilled human capital given by formula (7) and adjusted for the net productivity augmentation $\theta_t$ will exceed his wage based on the unskilled human capital obtained at the first stage of education according to its production function (6).

The kinked form of the college human capital production function (7), combined with pre-college preparation given by (6), implies that individual’s advanced human capital attainment exhibits increasing returns to ability, for which the quality of basic publicly provided education is a complementary input. This allows us to capture an important and arguably realistic property of the “ability premium” of college education: the skill upgrade that it provides to a highly able student is disproportionately larger than the one gained by a less able peer. Furthermore, while higher quality of public basic education “lifts all boats”, more able students will derive disproportionately greater benefits from it. This “ability premium” argument is used in some of the recent literature.
to explain the evidence of increasing dispersion of earnings. For example, Huggett et al. (2006) use life-cycle framework with a multi-stage Ben-Porath type model of human capital accumulation, which exhibits increasing returns to ability at higher stages of education, to explain the evolution of wage dispersion in the U.S. Another strand of models represented by Galor and Zeira (1993) is able to explain intergenerational persistence of inequality by the presence of credit constraints. The underlying mechanism, however, is fundamentally similar: the consequence of the borrowing constraints is that investment in education exhibits increasing returns to agents’ endowments (within a certain range). Restuccia and Urrutia (2004) use a calibrated model which includes explicit early and college stages of education to apportion the factors, including individual ability and borrowing constraints, responsible for the intergenerational persistence of income inequality. By contrast, in most models of public education, such as by Glomm and Ravikumar (1992), human capital accumulation exhibits decreasing returns to private inputs, which leads to vanishing relative variation of income.

2.4. Quality of Basic Education

We shall now introduce the per student basic education quality $E_t$, i.e. the public input in the basic education production function (6) provided in period $t$, as a function of the quality and quantity of teachers chosen by a government agency. Recall that only college educated individuals are eligible to be employed as teachers. Let $\Sigma_t$ be the set of individuals $\omega$ in generation $t$ employed as teachers. Let $z_t$ be the total number of teachers. Since population size was normalized to 1 in all generations, $z_t$ is also the fraction of teachers in the overall population in generation $t$, as well as the student-teacher ratio for generation $t + 1$ students.

We define the aggregate teacher quality as the aggregate human capital of teachers

---

4 Heckman and Cunha (2007) and related recent work appear to provide a unified framework for these approaches.
Likewise, the average teacher quality is given by \( \int_{\omega \in \Sigma_i} h_i^t(\omega) d \mu_i(\omega) \). The explicit account for the heterogeneity of teachers’ human capital attainment reflected in this formula is obviously an essential element of our model. Earlier papers, such as by Eckstein and Zilcha (1994), which explicitly modeled teacher human capital as an input in (compulsory) schooling, assumed that it equals to the average human capital of their generation.\(^5\)

We now define the per student quality of basic education as a Cobb-Douglas function of the quantity and aggregate quality of teachers:

\[
E_t = z_i^\gamma q_i^\nu
\]  

Note that this formula corresponds to the one used by Tamura (2001) who assumed in particular that the role of personal instruction, i.e. that of teacher-student ratio, is more important for schooling effectiveness than the average quality of teachers; in our formulation this means that \( \gamma \geq \nu \).

The special case of (10), when \( \gamma = \nu = 1 \), i.e.

\[
E_t = z_i \int_{\omega \in \Sigma_i} h_i^t(\omega) d \mu_i(\omega)
\]  

has a particularly straightforward interpretation. Assume that all teachers are perfectly sorted across classes, each class of equal size \( z_i^{-1} \), so that each is exposed student through his classes to a cross-section of teachers which perfectly represents their distribution of quality. Then the expression (11) which is equivalent to \( E_t = \int_{\omega \in \Sigma_i} h_i^t(\omega) d \mu_i(\omega) / z_i^{-1} \) can be interpreted as per student average teacher quality.

\(^5\) Hatzor (2008) contrasts such regime where teachers are selected at random from the population with the one where the quality of teachers is an optimal policy decision with a trade-off against their quantity. She focuses on comparing the implications of these regimes for growth and welfare in the framework of strategic interaction between the education and budgeting authorities of the government.
2.5. Government

The government funds and administers public education at the basic level with the goal of maximizing its quality $E_r$, as defined above, subject to the budget constraint given by the revenue from a uniform labor income tax at a flat rate $\tau_r$. To this end in each period $t$, the government must set teacher salary $I_t^{h}$ and the number of teachers to be hired $z_t$. As discussed in the Introduction, we postulate that all teachers in generation $G_t$ receive equal salary, according to a collective bargaining agreement. Since college educated individuals have an option to work in the production sector for a competitive wage as defined by the expression (5), the government’s choice of teacher salary $I_t^{h}$ will uniquely determine the highest level of human capital attainment $h_t$ among individuals who will choose to become teachers. Indeed it should satisfy the equation

$$\theta_t w_t h_t = I_t^{h} \quad (12)$$

Thus all college graduates with human capital level $h_t^*(\omega)$ at or below $h_t$ will be obviously motivated to accept employment as a teacher rather than work in the production sector. However, the government’s goal to maximize the overall education quality for a set number of teachers $z_t$ implies that the set $\Sigma_t$ of teachers the government will hire consists of all individuals whose level of human capital $h_t^*(\omega)$ attained in college falls into the interval $[h_t, \bar{h}]$ where the minimum teacher qualification threshold $h_t$ is determined by the intended number of teachers, i.e. the measure

$$z_t = \mu \left( \omega | h_t \leq h_t^*(\omega) \leq \bar{h} \right) \quad (13)$$

where the top cut-off $\bar{h}$ is determined, according to (12), by the teacher salary $I_t^{h}$ set by the government.

---

6 Since one’s work career is summarily represented in our model by one time period, we do not model the wage dynamics over the course of a worker’s or teacher’s career as he accumulates seniority and experience. The appropriate understanding of the income variables in this framework is that they represent aggregates over the entire career, such as respective present values at the career’s outset. While teachers’ union collective bargaining agreements stipulate wage differentials based on seniority, equation (12) should be understood as the comparison of respective aggregates over the course of the alternative careers in question.
Recalling the production functions of basic and advanced education given, respectively, by the expressions (6) and (7), we define the cut-off innate ability levels \( a_t \) and \( \bar{a}_t \) which characterize the teachers who possess, respectively, the cut-off levels of human capital \( h_t \) and \( \bar{h}_t \) induced by the government policy choice. In other words,

\[
\begin{align*}
a_t &= \frac{h_t + Bh_t^*}{(b + B)CE_{i-1}} \\
\bar{a}_t &= \frac{\bar{h}_t + Bh_t^*}{(b + B)CE_{i-1}}
\end{align*}
\]

For the government policy choice of \( I_t^h \), \( z_t \), to be feasible, the minimum teacher qualification threshold \( h_t \) defined by (13) obviously must belong the range of human capital levels attained by college graduates. In other words, the corresponding ability level \( a_t \) must exceed the college attendance cut-off level \( a_t^* \).

Thus according to (10) the government’s basic education quality optimization problem can be restated as

\[
\begin{align*}
\max_{z_t, \bar{h}_t} & \quad E_t \\
\text{subject to} & \quad (13) \\
& \quad z_t \theta_t w_t \bar{h}_t = T_t \quad \text{and} \quad a_t \geq a_t^* 
\end{align*}
\]

where \( T_t \) is the tax revenue collected by the government in period \( t \).

Thanks to our assumption of the uniform distribution of innate ability on the interval \([a, A]\) and according to the basic and advanced education production functions (6) and (7) we can simplify expressions (10) and (13), respectively, as

\[
\begin{align*}
E_t &= z_t^\prime q_t^\prime = \frac{z_t^\prime \left[ \bar{h}_t^2 - h_t^2 \right]}{2(A-a)(b+B)CE_{i-1}} \\
z_t &= \frac{\bar{h}_t - h_t}{(A-a)(b+B)CE_{i-1}}
\end{align*}
\]

and therefore problem (15) to maximize the quality of basic education \( E_t \) subject to the government budget constraint can be restated as
\[
\max_{z_i, \delta_i} \frac{z_i^r \left( \bar{h}_i^2 - \bar{h}_i^2 \right)^v}{2(A-a)(b+B)CE_{t-1}}^v
\]
subject to (17),

\[z_i \theta_i w_i \bar{h}_i = T_i \quad \text{and} \quad a_i \geq a_i^*\]
or equivalently, according to (17), as

\[
\max_{z_i, \delta_i} 2^{-v} \left( \bar{h}_i + \bar{h}_i \right)^v z_i^{r + v}
\]
subject to (17),

\[z_i \theta_i w_i \bar{h}_i = T_i \quad \text{and} \quad a_i \geq a_i^*\]

Note that the optimal lower and upper cut-off levels of teachers’ human capital \( h_i, \bar{h}_i \) are related through the optimal choice of their number \( z_i \) according to equation (17). The optimization in problem (18) thus expresses the trade-off between the quantity and quality of teachers to be hired. The quality of the top teacher \( \bar{h}_i \) will not only determine his salary \( I_i^h = \theta_i w_i \bar{h}_i \) due to his outside option as a skilled worker, but will set the identical salary for all other teachers according to the equal pay based collective bargaining agreement. Conversely, teacher salary \( I_i^h \) set by the government will uniquely determine the top teacher quality \( \bar{h}_i \). Therefore the total teachers’ wage bill in the government budget constraint is given by \( z_i \theta_i w_i \bar{h}_i \).

According to the relationships (14), the expression (17) is equivalent to

\[z_i = \frac{a_i - a_i^*}{A - a}\] (19)

Therefore using relationships (14) to express \( \bar{h}_i \) and \( \bar{h}_i \) and then eliminate \( a_i \) according to formula (19), we obtain

\[q_i = \frac{\bar{h}_i^2 - \bar{h}_i^2}{2(A-a)(b+B)CE_{t-1}} \left[ \left( b + B \right) CE_{t-1} a_i - Bh^* - \frac{1}{2} z_i (A-a)(b+B) CE_{t-1} \right] z_i\]
so we can restate the government’s education quality optimization problem (18) as

$$\max_{z_t, \pi_t} E_t = 2^{-\nu} \left[ 2(b + B)CE_{t-1}\pi_t - 2Bh^* - z_t(A - a) - (b + B)CE_{t-1} \right]^{\nu} z_t^{\nu}$$

subject to

$$z_t, \theta_t w_t \left[ (b + B)CE_{t-1}\pi_t - Bh^* \right] = T_t$$

$$\pi_t - z_t(A - a) \geq a_t^*$$

(20)

Figure 1 below offers an illustration for the government's education quality optimization problem in period $t$. The horizontal axis maps the ability of current workers. The vertical access represents wage income. For unskilled workers it is given by $w/h^*$, while for the skilled production sector workers it is $w/h^t$. The sloped line represents the distribution of incomes of production sector workers in generation $t$, in relation to individual ability according to the formulas (6) and (7), where the kink corresponds to the ability cut-off for college attendance $a_t^*$. 

![Figure 1: Basic Education Optimization Problem](image)
The vertical bars in the figure illustrate some of the feasible government education policy combinations of the number of teachers to be hired and teacher salaries: $(z^1_t, I^h_t), (z^2_t, I^h_t), (z^3_t, I^h_t)$, from left to right, satisfying the government’s budget constraint. The width of the bar corresponds to the number of teachers $z_t$ while its height stands for teacher salary $I^h_t$, so the bar’s area is the total expenditure. Policy $(z^1_t, I^h_t)$ represented by the far left bar in the figure is characterized by the lowest student-teacher ratio (the largest quantity of teachers) but also the lowest teacher quality. This policy option is highly inefficient in terms of education quality derived from the given budget revenue, as shown by the large deadweight loss represented by the triangular portion of the bar located above the sloped line: this is the excess salary paid to teachers above their opportunity cost value. The deadweight loss is the highest here because this policy involves hiring an inefficiently large number of low quality teachers who nevertheless have to be paid at the same rate as their best peer. The policy $(z^3_t, I^h_t)$ (far right in the figure) has the highest teacher quality. This set of education policies minimizes the dead weight loss associated with collective bargaining but still has low per pupil education quality due to small number of teachers, i.e. high student-teacher ratio. The interior policy option $(z^2_t, I^h_t)$ is where the per student education quality is maximized, providing the balance in terms of student-teacher ratio and the deadweight loss associated with the span of teacher quality.

3. General Equilibrium and Optimal Policy

We can now summarize the fundamental elements of the model and their relationships in a general equilibrium framework. We will first define the dynamic general equilibrium for given government education policy parameters and then incorporate the government’s optimal policy into the recursive dynamic general equilibrium framework.

Given the sequence of tax rates $\{r_t\}_{t=0}^\infty$ and the sequence of government education policy parameters $\{I^h_t, z_t\}_{t=0}^\infty$, i.e. teacher salaries and the numbers of teachers hired in
each period, respectively, as well as the initial period $t = 0$ aggregate supply of capital $K_0$, the distributions of the retirees’ consumption levels $c_{t,0}(\omega)$, and per student basic education quality $E_{t-1}$ provided to generation $G_0$ individuals as children, we define the **dynamic general equilibrium** as a collection of sequences of

(a) factor prices $\{(1+r_{t+1}), w_t, \theta_t w_t\}_{t=0}^\infty$ respectively of physical, unskilled and skilled human capital as inputs in production in period $t$;

(b) aggregate variables $\{Y_t, K_t, H^u_t, H^v_t, T_t, E_t, a_t^s\}_{t=0}^\infty$, i.e., respectively, aggregate output, inputs of physical, unskilled and skilled human capital in production, government’s tax revenue, the quality of basic education provided to each student in period $t$, as well as the endogenous innate ability cut-off for college attendance;

(c) distributions of individual consumption and education decisions $\{c_{t,t}(\omega), c_{t,t+1}(\omega), h^u_t(\omega), h^v_t(\omega)\}_{t=0}^\infty$, as well as employment decisions by college graduates such that

(i) the factor prices are determined competitively, i.e. are set equal to the marginal products of respective inputs:

$$1 + r_{t+1} = \alpha D K^\alpha_t \left[ H^u_t + \theta_t H^v_t \right]^{-1-\alpha}, \quad w_t = (1-\alpha) D K^\alpha_t \left[ H^u_t + \theta_t H^v_t \right]^{-\alpha}$$

(ii) each individual $\omega \in [0,1]$ in generation $G_t$ makes a decision whether to go to college and if so whether to be employed as a teacher or in the production sector so as to maximize his income while taking as given basic education quality $E_{t-1}$, production sector wage rates $w_t$ and $\theta_t w_t$ (for unskilled and skilled labor, respectively), teacher salary $I^h_t$ and the number of teachers $z_t$ to be hired, whereas his human capital level $h^u_t(\omega)$ or $h^v_t(\omega)$ (depending on his college attendance decision) is determined according to the education production functions (6) and (7); (note that according to equation (12) and the collective bargaining agreement, teacher salary will exceed production sector wage for all
but the top quality teacher, so the government teacher employment limit $z_t$ will bind;

(iii) based on his income $I_t(\omega)$ determined according to (ii), each individual $\omega \in [0,1]$ in generation $G_t$ makes his young- and old-age consumption decisions $c_{t,\omega}(\omega), c_{t+1,\omega}(\omega)$ by solving the optimization problem (2)- (3) while taking the rates of interest $1 + r_{t+1}$ and tax $\tau_t$ as given;

(iv) the quality of basic education $E_t$ provided to generation $G_{t+1}$ individuals (as children) is determined by the expression (10) while the set of teachers $\Sigma_t$ is defined by individual employment decisions according to (ii) and the number of teachers hired $z_t$ is as given by the government’s policy;

(v) the markets for goods, physical capital, and skilled and unskilled labor clear in each period:

$$Y_t = \int_{\omega \in [0,1]} c_{t,\omega}(\omega) d\mu_t(\omega) + \int_{\omega \in [0,1]} c_{t-1,\omega}(\omega) d\mu_{t-1}(\omega),$$

$$K_t = (1 + r_{t+1})^{-1} \int_{\omega \in [0,1]} c_{t,r+1}(\omega) d\mu_t(\omega),$$

$$H_t^u = \int_{\omega \leq a(\omega) \leq a^*_t} h_t^u(\omega) d\mu_t(\omega),$$

$$H_t^{sv} = \int_{a^*_t \leq a(\omega) \leq A} h_t^i(\omega) d\mu_t(\omega) - \int_{\omega \in \Sigma_t} h_t^i(\omega) d\mu_t(\omega),$$

where the ability cut-off for college attendance $a^*_t$ is determined by individual college attendance decisions as defined in (ii);

(vi) the aggregate tax revenue is composed of labor income taxes collected from all categories of employees, i.e.

$$T_t = \tau_t \left( w_t H_t^u + \theta_t w_t H_t^{sv} + z_t I_t^b \right)$$

We can now define the government’s optimal education policy in period $t$ recursively, based on the above general equilibrium construct. Namely, the government
chooses teacher salaries $I^h_t$ and the number of teachers $z_t$ for period $t$ by solving the optimization problem (18) (or, equivalently, problem (20)) where the top teacher quality $\bar{h}_t$ is determined by equation (12), while taking as given the economy’s general equilibrium values of production sector wage rate $w_t$, aggregate tax revenue $T_t$ and the distribution of skilled human capital attainment $h'_t(\omega)$ by generation $G_t$ individuals.

Noting the mutual dependence of the general equilibrium variables in period $t$ and the government’s optimal education policy we define the Education-Economy recursive dynamic equilibrium (RDE for brevity) as a fixed point of this relationship, recursively determined for each period $t$.

Remark. Since we assumed that individuals make a decision whether to attend college solely on the basis of maximizing income, it is clear that the ability cut-off for college attendance $a^*_t$ defined in part (ii) of the definition of the dynamic general equilibrium, should satisfy inequality

$$a^*_t \leq \frac{1}{CE_{t-1}} \frac{\theta Bh^*_t}{\theta (b+B)-1}$$

(26)

Indeed, according to (6), (7) and (4), (5), an individual with ability exceeding the right hand side of (26) will certainly increase his income by going to college. In fact, we will show in the next section that in the RDE inequality (26) is satisfied as equality, i.e. formula (9) is true.

4. Analysis of the Model

To simplify the exposition we will assume henceforth without any substantive loss of generality that parameter $a = 0$, thus innate ability in each generation is distributed uniformly on the interval $[0, A]$. We begin by analyzing the government’s optimal education policy problem given by the equivalent formulations (18) or (20).
We impose the following restrictions on the economy’s parameters, where $E_{-1}$ is an exogenously given per student basic education quality provided to generation $G_0$ individuals.

**Assumption 1.**

\[
\left( \frac{\nu}{\gamma} (b + B) AC \left( 1 - \tau_t \right) \right)^{1/2} \left( \frac{\gamma}{2\nu + \gamma} \right) \left( 1 - \tau_t \right) \left( 1 + \frac{Bh^*}{(b + B) ACE_{-1}} \right) > 1
\]

is true for any $t = 0, 1, ..., $ for any $t = 0, 1, ..., $.

**Assumption 2.**

\[
\left( \frac{\nu(1 - \tau_t) - 1}{\gamma} \right) \left( \frac{\gamma}{2\nu + \gamma} \right) \left( 1 - \tau_t \right) \left( 1 - \frac{Bh^*}{(b + B) ACE_{-1}} \right) > \frac{1}{\theta_t (b + B)}
\]

is true for any $t = 0, 1, ..., $.

The above assumptions require that education taxes $\tau_t$ not be too small while not exceeding $1 - \frac{\gamma}{2\nu}$, which of course imposes a requirement that $\gamma$, the relative importance of the teacher-student ratio for schooling effectiveness (see formula (10)), should not be substantially greater than $\nu$, the relative importance of the teacher quality. The main thrust of the assumptions, however, concerns the parameters which characterize educational gains. Assumption 1 is satisfied if parameter $C$ characterizing the human capital gains in basic education according to (6), is sufficiently large. Assumption 2 will hold if $(b + B)$, a productivity characteristic of the college education production function (7), is large enough.

Based on these assumptions we will characterize the optimal solution of the education quality optimization problem in terms of the optimal number of teachers $z_t$ for period $t$, the corresponding range of teachers’ human capital, i.e. its upper and lower cut-off values $\bar{h}_t$, $\underline{h}_t$ induced by the policy and the corresponding cut-off innate ability levels
In the process we will establish the following important facts (see Appendix 1 for the proofs):

Lemma 1 (Growth of Basic Education Quality). The recursive equilibrium dynamics exhibits sustained growth of the quality of per student basic education. Specifically, there is a rate $g > 1$ such that $E_t \geq gE_{t-1}$ is true for all $t = 0, 1, \ldots$

Lemma 2 (The Interiority Property). In the recursive dynamic equilibrium, the ability of the least qualified teacher exceeds the college attendance cut-off ability in all time periods, i.e. $a_t > a_t^*$ is true for $t = 0, 1, \ldots$. Thereby the human capital of the least qualified teacher will not be the lowest among his contemporary college graduates.

Lemma 3. The ability cut-off for college attendance $a_t^*$ satisfies equality (9), i.e.

$$a_t^* = \frac{1}{CE_t} \theta_t B h_t^*$$

which means that an individual will choose to attend college if and only if his resulting skilled human capital given by formula (7) adjusted for the net productivity augmentation $\theta_t$ will exceed his unskilled human capital derived from the basic stage of education according to its production function (6).

We now proceed to solving the optimization problem (20). According to the teacher salary equation (12) and the tax revenue formula (25), the government budget constraint can be stated as

$$(1-\tau_t)z_t \theta_t \bar{h}_t = \tau_t \left(H_t^\alpha + \theta_t H_t^\gamma\right)$$

(27)

Using the education production functions (6) and (7) and the assumption that innate ability is uniformly distributed on $[0, A]$ we can rewrite the general equilibrium relationships (23), (24) as
\[ H_i^a = CE_{i-1} \int_0^1 \frac{A}{A} \, da = \frac{(a_i^*)^2}{2A} \, CE_{i-1} \quad (28) \]

\[ H_i^{\nu} = \int_{a_i^*}^A \left[ (b + B) CE_{i-1} (a - Bh^{\nu}) \right] \frac{1}{A} \, da - \int_{\bar{a}}^A \left[ (b + B) CE_{i-1} (a - Bh^{\nu}) \right] \frac{1}{A} \, da \]

\[ = \frac{(b + B) CE_{i-1}}{2A} \left[ A^2 - (a_i^*)^2 - (\bar{a}_i)^2 + (a_i^*)^2 \right] - \frac{Bh^{\nu}}{A} \left[ A - a_i^* - \bar{a}_i + a_i \right] \quad (29) \]

Therefore expressing \( \bar{h}_r \) through \( \bar{a}_r \) according to the relationship in (14) we can rewrite the budget constraint (27) as

\[ (1 - \tau_i) \theta_i \bar{z}_i \left( (b + B) CE_{i-1} \bar{a}_i - Bh^{\nu} \right) = \]

\[ \frac{\tau_i CE_{i-1}}{2A} \left[ (a_i^*)^2 + \theta_i (b + B) (A^2 - (a_i^*)^2 - (\bar{a}_i)^2 + (a_i^*)^2) \right] - \frac{\tau_i \theta_i Bh^{\nu}}{A} \left[ A - a_i^* - \bar{a}_i + a_i \right] \quad (30) \]

We now eliminate variables \( a_i^* \) and \( \bar{a}_i \) from (30) by substituting the value of \( a_i^* \) given by (9) according to Lemma 3, and using the expression \( \bar{a}_i = \bar{a}_i - Az_i \) which follows from the relationship (19) since we set \( a = 0 \). This immediately turns expression (30) into a linear equation in terms of variable \( \bar{a}_i \) which yields

\[ \bar{a}_i = \frac{z_i \tau_i A}{2} + \frac{Bh^{\nu}}{(b + B) CE_{i-1}} + \frac{\tau_i A}{2z_i} \left( 1 - \frac{2Bh^\nu}{A(b + B) CE_{i-1}} \right) \left( \frac{\theta_i (b + B - 1)(b + B) (ACE_{i-1})^2}{\theta_i (b + B - 1)(b + B) (ACE_{i-1})^2} \right) \quad (31) \]

This expression incorporates the government budget constraint of problem (20). That problem’s objective function, upon substituting the expression (31) for \( \bar{a}_i \), becomes a function of a single variable \( z_i \). We will solve for its unconstrained maximization and then refer to Lemma 2 which ensures that the only remaining constraint \( \bar{a}_i - Az_i \geq a_i^* \) in the government optimization problem (20) is automatically fulfilled.

Thus we are looking at the unconstrained maximization of the following function:

\[ F(z_i) = q_i^\nu \bar{z}_i^\nu \]

\[ = \left( \frac{\tau_i (b + B) ACE_{i-1}}{2} - \tau_i Bh^{\nu} + \frac{\tau_i Bh^{\nu} \bar{z}_i^\nu}{2} \left( 1 - \frac{\tau_i (b + B) ACE_{i-1}}{2} \right) \right) \bar{z}_i^\nu \]

Its first order necessary condition is given by the equation
\[
\gamma z_t^{\gamma-1} q_t^* - \nu (1 - \tau_t) (b + B) ACE_{t-1} z_t^{\tau_{t-1}} = 0
\]  

yielding unique non-negative solution:

\[
z_t = \left( \frac{\gamma}{2\nu + \gamma} \left( \frac{\tau_t}{1 - \tau_t} \right) \right)^{1/2} \left( 1 - \frac{2Bh^*}{(b + B)ACE_{t-1}} + \frac{\theta (Bh^*)^2}{(\theta (b + B) - 1)(b + B)(ACE_{t-1})^2} \right)^{1/2}
\]  

It is straightforward to verify that this solution also satisfies the second order sufficient condition of the maximization problem. Substituting expression (33) back into formula (31) we obtain

\[
\bar{a}_t = \frac{\tau_t A z_t}{2} + \frac{Bh^*}{(b + B)CE_{t-1}} + \frac{\tau A}{2z_t} \left( 1 - \frac{2Bh^*}{(b + B)ACE_{t-1}} + \frac{\theta (Bh^*)^2}{(\theta (b + B) - 1)(b + B)(ACE_{t-1})^2} \right)
\]

which simplifies, by using equation (33) again, into

\[
\bar{a}_t = \frac{\tau_t A z_t}{2} + \frac{Bh^*}{(b + B)CE_{t-1}} + \frac{2\nu + \gamma}{2\gamma} (1 - \tau_t) A z_t = \frac{Bh^*}{(b + B)CE_{t-1}} + A z_t \left( \frac{\nu (1 - \tau_t)}{\gamma} + \frac{1}{2} \right)
\]

Recall that \( a = \bar{a}_t - A z_t \) according to (19) since we set \( a = 0 \). Applying this to (34) we obtain

\[
a_t = \frac{Bh^*}{(b + B)CE_{t-1}} + A z_t \left( \frac{\nu (1 - \tau_t)}{\gamma} - \frac{1}{2} \right)
\]

Observe that the education policy optimization as well as the individuals’ and the production sector’s general equilibrium reactions are determined recursively. Indeed, according to expressions (33) - (35), education quality \( E_{t-1} \) uniquely determines optimal education policy in period \( t \), i.e. the number of teachers, as well as the range of their innate abilities and thereby, due to (14), the range of their human capital attainment. This in turn will uniquely determine college attendance and employment decisions by generation \( t \) individuals, hence their incomes and their allocations. Government’s education policy will also determine the current period’s basic education quality \( E_t \), so the recursion continues.
Consider now the effect of the previous period’s education quality $E_{t-1}$ on education decision variables in period $t$. By differentiating the expressions (9) and (33) we obtain:

$$\frac{\partial a^*_t}{\partial E_{t-1}} = \frac{-\theta_t B^*_t}{\left(\theta_t (b + B) - 1 \right) CE_{t-1}^2} < 0$$  \hspace{1cm} (36)$$

$$\frac{\partial z_t}{\partial E_{t-1}} = \frac{1}{z_t} \left( \frac{\gamma}{2v + \gamma} \right) \left( \frac{1}{1 - \tau_t} \right) \left( \frac{B^*_t}{(b + B) ACE_{t-1}^2} \right) \left( 1 - \frac{a^*_t}{A} \right) > 0$$ \hspace{1cm} (37)$$

According to (34) and (35), respectively, we can write

$$\frac{\partial a_t}{\partial E_{t-1}} = -\frac{B^*_t}{(b + B) CE_{t-1}^2} + \frac{A}{z_t} \left( \frac{2\nu (1 - \tau_t) + \gamma}{2\gamma} \right) \left( \frac{\gamma}{2v + \gamma} \right) \left( \frac{1}{1 - \tau_t} \right) \left( \frac{B^*_t}{(b + B) ACE_{t-1}^2} \right) \left( 1 - \frac{a^*_t}{A} \right)$$ \hspace{1cm} (38)$$

$$\frac{\partial z_t}{\partial E_{t-1}} = \frac{B^*_t}{(b + B) CE_{t-1}^2} \left( -1 + z_t^{-1} \left( \frac{2\nu (1 - \tau_t) + \gamma}{2\gamma} \right) \left( \frac{\gamma}{2v + \gamma} \right) \left( \frac{1}{1 - \tau_t} \right) \left( 1 - \frac{a^*_t}{A} \right) \right)$$ \hspace{1cm} (39)$$

Note that since $\frac{\theta_t (b + B)}{\theta_t (b + B) - 1} > 1$, the following inequality is true

$$1 - \frac{2B^*_t}{(b + B) ACE_{t-1}} + \frac{\theta_t \left( B^*_t \right)^2}{(b + B) \left( \theta_t (b + B) - 1 \right) \left( ACE_{t-1} \right)^2} > \left( 1 - \frac{B^*_t}{(b + B) ACE_{t-1}} \right)^2$$ \hspace{1cm} (40)$$

Therefore according to (33)

$$z_t > \left( \frac{\gamma}{2v + \gamma} \left( \frac{1}{1 - \tau_t} \right) \right)^{1/2} \left( 1 - \frac{B^*_t}{(b + B) ACE_{t-1}} \right) > \left( \frac{\gamma}{2v + \gamma} \left( \frac{1}{1 - \tau_t} \right) \right)^{1/2} \left( 1 - \frac{\theta_t B^*_t}{\theta_t (b + B) - 1 \left( ACE_{t-1} \right)^2} \right) = \left( \frac{\gamma}{2v + \gamma} \left( \frac{1}{1 - \tau_t} \right) \right)^{1/2} \left( 1 - \frac{a^*_t}{A} \right)$$
Thus the expression (38) will be negative as long as the inequality
\[
\left( \frac{\gamma}{2v + \gamma} \frac{\tau_i}{1 - \tau_i} \right)^{1/2} > \left( \frac{\gamma}{2v + \gamma} \frac{\tau_i}{1 - \tau_i} \right)
\]
is true, which is certainly the case under the non-binding condition \( \tau_i < \frac{4}{5 + 2v/\gamma} \) on the tax rate (see Appendix 1 for the proof of this assertion). Comparing expressions (38) and (39) one can see that negativity of (38) implies the same for (39). Therefore we can conclude that
\[
\frac{\partial \bar{a}_i}{\partial E_{t-1}} < 0, \quad \frac{\partial \underline{a}_i}{\partial E_{t-1}} < 0
\]
Combining these facts with Lemma 1, which shows that education quality \( E_{t-1} \) does in fact grow over time, we obtain our central result.

**Theorem 1 (Dynamics of the Quantity and Quality of Teachers).** The recursive dynamic equilibrium (RDE) exhibits the following evolution of education policy variables:

- the quantity of teachers \( z_i \) grows over time;
- the relative quality of teachers characterized by the range of their innate abilities falls: both the upper and the lower cut-offs \( \bar{a}_i, \underline{a}_i \) decrease over time;
- the college attendance ability cut-off \( a^*_i \) also drops over time and (according to Lemma 2) remains consistently below the lower ability cut-off for teachers \( \underline{a}_i \).

Recall that according to relationships (14)
\[
\bar{h}_i = \bar{a}_i \left( b + B \right) CE_{t-1} - Bh^* \quad \text{and} \quad \underline{h}_i = \underline{a}_i \left( b + B \right) CE_{t-1} - Bh^*
\]
Therefore due to (34) and (35), respectively, as well as to (33) we can write
This leads to the following.

**Corollary.** As the relative quality of teachers falls over time in the RDE (according to the Theorem), the absolute quality of teachers characterized by their human capital attainment grows: both the human capital of the top teacher and the least qualified one, $\bar{h}_t$, $\underline{h}_t$, rise over time.

**Discussion.** The intuition for the above results derives from the mechanics of economic growth in our model. Rising per student quality of basic education (Lemma 1) opens up the opportunity to pursue higher education for an expanding group of students. Namely, college attendance becomes worthwhile for an ever broader population, adding on students with relatively low ability. At the same time, the human capital attainment of highly able students increases disproportionately relative to their less able peers due to increasing returns to ability exhibited by the college education production function (7). Thus economic growth drives the rise of income inequality within the group of college graduates. As a result, hiring high ability individuals as teachers becomes a relatively
more expensive option, which pushes the quality-quantity trade-off in the education policy in favor of the latter.

Figure 2 below illustrates the evolution of the optimal education policy choices along with the evolution of income distribution in successive time periods, $t$ and $t+1$. The part of the graph pertaining to period $t$ corresponds to the illustration given in Figure 1 of Section 3: the dashed sloped line represents the income distribution of production sector workers, while the light shaded bar depicts the optimal education policy (the bar’s width $z_t$ is the number of teachers while its height is teacher salary).

The solid sloped line represents wage income distribution in period $t+1$. It exhibits a kink further to the left, i.e. a lower college attendance ability cut-off than in period $t$, consistent with Theorem 1. Furthermore, the comparison of the solid and dashed lines confirms the fact discussed above that the benefit derived by skilled individuals from the quality of basic education grows disproportionately with their ability. This implies the rising relative cost of highly able teachers and results in the shift of the quantity-quality trade-off depicted by the position of the dark shaded bar, the
government optimal education policy choice in period \( t + 1 \), to the left of the period \( t \)’s light shaded bar, as stated in Theorem 1.

The argument in the above discussion concerning the growing relative cost of high ability individuals is made explicit by the following result which characterizes the evolution of income inequality in our model.

Based on the income formulas (4)-(5), the human capital accumulation formulas (6)-(7) and using the uniform distribution of abilities as well as the formula (9) for the threshold ability between the groups, we can obtain the mean income of unskilled individuals:

\[
T^u_t = \frac{I^u_t(a^*_t)}{2} = \frac{w_t \theta_t \theta_t B h^*_t}{2(\theta_t (b + B) - 1)} = \frac{w_t C E_{t-1} a^*_t}{2}
\]

and the mean income of the skilled (ignoring the distortion due to collective bargaining in the education sector):

\[
T^s_t = \frac{I^s_t(a^*_t) + I^s_t(A)}{2} = \frac{w_t \theta_t \left( (b + B) \theta_t B h^*_t \right)}{2 \left( \theta_t (b + B) - 1 \right)} + A \left( b + B \right) C E_{t-1} - 2 B h^*_t
\]

Thus the inequality between the groups can be characterized by

\[
\sigma^s/u_t = \frac{T^s_t}{T^u_t} = \frac{A \left( b + B \right) C E_{t-1} \left( \theta_t (b + B) - 1 \right)}{B h^*_t} + \left( 2 - \theta_t (b + B) \right)
\]

This expression obviously increases in basic education quality, which according to Lemma 1 rises over time.

The inequality within the skilled group (ignoring the aforementioned distortion) is characterized by

\[
\sigma^s_t = \frac{I^s_t(A)}{I^s_t(a^*_t)} = \frac{(b + B) A C E_{t-1} - B h^*_t}{(b + B) a^*_t C E_{t-1} - B h^*_t} = \left( \theta_t (b + B) - 1 \right) \left( b + B \right) A C E_{t-1} - B h^*_t
\]

which also grows with the rise of basic education quality. One can rewrite this expression, according to (9), as

\[
\sigma^s_t = \left( \frac{B h^*_t}{C E_{t-1}} \right) \left( A - \frac{B h^*_t}{C E_{t-1}} \right)
\]
This demonstrates that the upward trend of the absolute disparity between the highest and lowest incomes of skilled workers can be attributed to two factors: (i) the rise of basic education quality $E_{t-1}$ which increases the human capital and thereby the incomes of all workers, but disproportionately more so at the high end of the ability distribution; (ii) falling, according to Lemma 2, college attendance ability cut-off $a_t^*$, which brings lower ability workers into the fold of the skilled hence increasing the intra-group inequality. We summarize the above results as Theorem 2.

**Theorem 2 (The Evolution of Income Inequality).** The recursive equilibrium dynamics exhibits growing inequality within the group of skilled individuals, as well the rise in inequality between this group and the unskilled.

As we discussed above and in Section 2, this result is due to the absolute growth of the quality of per student publicly provided basic education, which has unequal impact on individuals across the distribution of abilities because of the complementary relationship between individual ability and quality of education.\(^7\)

The recent growth literature which presents evidence of rising dispersion of incomes of skilled workers over the last decades attributes this to the skill biased nature of technological change (see Acemoglu (1998), (2000), Galor and Moav (2000)). While our results on income inequality have been derived exclusively from the impact of growing public provision of basic education, one should expect that a rise in skill premium $\theta_t$ will exacerbate these effects.

We will now introduce the exogenous skill biased technological change into the model, given by positive shocks to the skill premium coefficients $\theta_t$, and will explore its effects on the education policy variables and the quality of education. Specifically, we consider the recursive dynamic equilibrium (RDE) corresponding to the original

---

\(^7\) A somewhat similar argument for the magnifying effect that greater public education funding may have on income inequality is advanced by Glomm and Kaganovich (2003) in the presence of complementarity between public and parental private inputs, imperfect altruism, and borrowing constraints. The fact of such complementary relationship and its implications for inequality was documented for the case of Britain by LeGrand (1982).
(benchmark) exogenously given sequence \( \{\theta_t\}_{t=0}^\infty \) and assume that the productivity augmentation of skilled labor receives a positive shock from time \( t_0 \) on, i.e. that for \( t = t_0, t_0+1, \ldots \) the values \( \theta_t \) are replaced with some \( \theta_t' > \theta_t \). We will characterize the effect of this exogenous change on the RDE, particularly on the education policy variables. We obtain the following comparative dynamics result (see Appendix for the proof).

**Theorem 3 (The Effect of Skill Biased Technological Change).** Consider the comparative dynamics experiment described above where skill premium coefficients \( \theta_t \) receive a positive shock from period \( t_0 \) on. The corresponding recursive dynamic equilibrium, relative to the benchmark RDE, will be characterized, for \( t \geq t_0 \), by

- lower quantity of teachers \( z_t \);
- lower aggregate quality of teachers \( q_t \) and therefore
- lower quality of basic education \( E_t \).

Note the negative effect on both the number and aggregate quality of teachers which is due to an upward shock to the cost of skilled labor. The Theorem thus shows that the technological change biased toward skilled labor will have a detrimental effect on the absolute quality of basic education, exacerbating the negative effect of a secular downward trend in the relative quality of teachers stated in Theorem 1. These results will apply, in particular, when the education tax rate \( \tau \) stays constant, which means that education budget grows at the rate of GDP growth. This leads to an important implication of our result: given the negative impact of rising skill premium on the quality of education, a policy aimed at neutralizing this effect would require an increase in funding of education (assuming no change in the institutional setup of the school system and teachers’ labor market) at a rate faster than GDP growth, i.e. raising the fraction of GDP devoted to education. A proper interpretation of this conclusion follows directly from our results: the relative cost of maintaining a given teacher quality standard will rise over time.
5. Conclusions

Over the last forty years, education policy in the U.S. has changed significantly, focusing in particular on lowering the student-teacher ratio. We have developed a model which offers an insight into this evolution by relating it to the changes in the US economy characterized by rising skill premium and overall income inequality. We argue that teacher wage compression due in large part to collective bargaining agreements has a significant effect on decisions concerning quantity-quality trade-offs in hiring teachers. Our model predicts that as incomes of college educated individuals rise and become more dispersed, education policy-makers are forced to adjust relative teacher salaries and thereby quality standards. Education quality is optimized by lowering relative quality of teachers while increasing their numbers. This causes the higher ability college educated people to choose private sector employment which offers higher reward to skilled workers.

We argue moreover that a rise in skill premium caused, in particular, by skill biased technological change will exacerbate the negative trends in the relative quality of education. Indeed, the labor of college graduates will further appreciate relative to the average wage and hence relative to the tax revenue. Countering this trend would therefore require an increase in the share of GDP spent on basic education, assuming that the institutional setup of the school system remains unchanged.

Our finding that skill biased technological change can have a negative effect on the quality of education is an interesting case of negative feedback, since SBTC literature points to the rise in the supply of skill due to growing availability of education as its underlying cause. Furthermore, this leads to an issue which appears important for future research on the aggregate long term effects of SBTC: as the technical change brings about productivity gains, one needs to factor in its effects on the cost and quality of education and the corresponding policy responses in order to assess the full long-term impact.
References


Appendix 1. Proofs

We will first prove Lemmas 1 and 2 under the hypothesis that Lemma 3 is correct. We will then prove that Lemma 3 is indeed correct in the recursive dynamic equilibrium, and thereby the imposition of the hypothesis will not have diminished the generality of (or create circularity problems with) the argument.

Proof of Lemma 1.

Recall that according to (20) \( E_t = \left( z_t (b + B) CE_{t-1} \alpha_t - z_t Bh^* - \frac{1}{2} (b + B) ACE_{t-1} \tau_t^2 \right)^{\nu} \).

Substituting the expression for \( \bar{a}_t \) given in (34), we obtain

\[
E_t = \left( A(b + B) CE_{t-1} \left( \frac{\nu (1 - \tau_t)}{\gamma} \right) \right)^{\nu} z_t^{2 \nu + \gamma}, \text{ or according to (33)}
\]

\[
E_t = \left( \frac{A(b + B) CE_{t-1} \nu (1 - \tau_t)}{\gamma} \right)^{\nu}.
\]

\[
\left( \frac{1}{2 \nu + \gamma} \right) \frac{\tau_t}{1 - \tau_t} \left( \frac{2 Bh^*}{(b + B) ACE_{t-1}} + \frac{\theta_t (Bh^*)^2}{(\theta_t (b + B) - 1)(b + B)(ACE_{t-1})^2} \right)^{1 + 2 \nu} > 1
\]

Note that since \( \frac{\theta_t (b + B)}{\theta_t (b + B) - 1} > 1 \), the following inequality is true

\[
1 - \frac{2 Bh^*}{(b + B) ACE_{t-1}} + \frac{\theta_t (Bh^*)^2}{(\theta_t (b + B) - 1)(b + B)(ACE_{t-1})^2} > \left[ 1 - \frac{Bh^*}{(b + B) ACE_{t-1}} \right]^2
\]

Therefore we can write

\[
E_t > \left( \frac{\nu}{\gamma} (b + B) ACE_{t-1} (1 - \tau_t) \right)^{\nu} \left( \frac{1}{2 \nu + \gamma} \right) \left( \frac{\tau_t}{1 - \tau_t} \right)^{1 + 2 \nu} \left( 1 - \frac{Bh^*}{(b + B) ACE_{t-1}} \right)^{2 \nu + \gamma}
\]

Thus, in order to prove the Lemma it is sufficient to show that for all \( t = 0,1, \ldots \)

\[
\left( \frac{\nu}{\gamma} (b + B) AC (1 - \tau_t) \right)^{\frac{1}{1 + 2 \nu}} \left( \frac{1}{2 \nu + \gamma} \right)^{\frac{1}{1 + 2 \nu}} \left( \frac{\tau_t}{1 - \tau_t} \right)^{\frac{1}{1 + 2 \nu}} \left( 1 - \frac{Bh^*}{(b + B) ACE_{t-1}} \right)^{\gamma + \nu} > 1
\]

which is indeed true according to Assumption 1 and by the induction argument. 

37
Proof of Lemma 2.

Based on Lemma 3 we use expression (9) for \( t^* \). Then according to (35) our task of proving the inequality \( t^* > a^*_i \) is equivalent to verifying the inequality

\[
\frac{Bh^*}{(b+B)CE_{i-1}} + A\tau_i \left( \frac{\gamma(1-\tau_i)}{2} - \frac{1}{2} \right) > \frac{1}{CE_{i-1}} \frac{\theta_i Bh^*}{\theta_i (b+B) - 1} \text{ or } A\tau_i \left( \frac{\gamma(1-\tau_i)}{2} - \frac{1}{2} \right)
\]

Upon substituting the expression (33) for \( z_i \), the last inequality becomes

\[
A \left( \frac{\gamma(1-\tau_i)}{2} - \frac{1}{2} \right) \left( \frac{\gamma}{2\nu + \gamma} \right) \left( \frac{\tau_i}{1-\tau_i} \right)^{1/2}.
\]

\[
1 - \frac{2Bh^*}{(b+B)ACE_{i-1}} + \frac{\theta_i (Bh^*)^2}{(\theta_i (b+B) - 1)(b+B)(ACE_{i-1})^2}
\]

\[
\left( \frac{\gamma(1-\tau_i)}{2} - \frac{1}{2} \right) \left( \frac{\gamma}{2\nu + \gamma} \right) \left( \frac{\tau_i}{1-\tau_i} \right)^{1/2} \left( 1 - \frac{Bh^*}{(b+B)ACE_{i-1}} \right) > \frac{1}{\theta_i (b+B)}
\]

Under Lemma 3 the right hand side in (42) is smaller than \( \frac{A}{\theta_i (b+B)} \) since \( a^*_i < A \).

Therefore according to (41) in order to prove inequality (42) it is by far sufficient to establish

\[
\left( \frac{\gamma(1-\tau_i)}{2} - \frac{1}{2} \right) \left( \frac{\gamma}{2\nu + \gamma} \right) \left( \frac{\tau_i}{1-\tau_i} \right)^{1/2} \left( 1 - \frac{Bh^*}{(b+B)ACE_{i-1}} \right) > \frac{1}{\theta_i (b+B)}
\]

which is indeed true for all \( t = 0,1,... \) according to Assumption 2 combined with Lemma 1.

\footnote{\textcopyright}
Proof of Lemma 3.

The above proofs were based on the hypothesis that Lemma 3 is correct, i.e. that the ability cut-off for college attendance $a^*_t$ satisfies equality (9), i.e. we proved that if college attendance cut-off ability is $a^*_t = \frac{1}{CE_{t-1}} \frac{\theta, Bh^*}{\theta_t (b + B) - 1}$ then the optimal education policy requires that all teachers’ ability strictly exceed this threshold. This in turn means that the marginal college graduate will be employed in the production sector. As we explained after stating equality (9), if an individual with ability below $a^*_t$ attended college his skilled human capital adjusted for the net productivity augmentation $\theta_t$ will be inferior to his unskilled human capital derived from the first stage of education, therefore a job in production sector’s skilled labor force would not compel such individual to attend college. Thus the only way the violation of Lemma 3 could occur is if such individual had an opportunity to be hired as a teacher. Compare, however, optimization problem (20) where $a^*_i < \frac{1}{CE_{t-1}} \frac{\theta, Bh^*}{\theta_t (b + B) - 1}$ to the one with $a^*_i = \frac{1}{CE_{t-1}} \frac{\theta, Bh^*}{\theta_t (b + B) - 1}$.

One can easily see that the only difference would be lower tax revenue $T_i$ in the former case. Therefore such government policy would be inferior to the one where $a^*_i = \frac{1}{CE_{t-1}} \frac{\theta, Bh^*}{\theta_t (b + B) - 1}$. Thus the latter indeed characterizes the recursive dynamic equilibrium optimum, i.e. Lemma 3 is correct. $\Box$
Proof of Inequality (41).
We rewrite the inequality by squaring its both sides:
\[
\left(\frac{\tau_i}{2\nu + \gamma}\right) > \left(\frac{\tau_i}{1-\tau_i}\right) \left(\frac{2\nu(1-\tau_i) + \gamma}{2\gamma}\right)
\]
which can be simplified: \((1-\tau_i)^2 + 2\nu + 1 > \tau_i^2 + (1-\tau_i)^2 + \frac{1}{2}\). This last inequality is certainly true if \((1-\tau_i)^2 + 2\nu + 1 > \tau_i^2 + (1-\tau_i)^2 + \frac{1}{2}\) which reduces to \(\tau_i < \frac{4}{5+2\nu/\gamma}\). 3

Proof of Theorem 3.
The proof will proceed by induction argument.
Consider firstly the effect of a positive shock to coefficient \(\theta_i\) in period \(t=t_0\) on education policy variables in this same period. According to (33) direct derivative \(\frac{\partial z_i}{\partial \theta_i}\)
has the same sign as \(\frac{\partial}{\partial \theta_i} \left[ \frac{\theta_i}{(b+B)+1} \right] \), i.e. negative, so we can write
\[
\frac{\partial z_i}{\partial \theta_i} < 0 \quad (43)
\]
Therefore, according to (34) and (35), respectively, we can write
\[
\frac{\partial a_i}{\partial \theta_i} = \frac{\partial a_i}{\partial z_i} \frac{\partial z_i}{\partial \theta_i} = A(1.5 - \tau_i) \frac{\partial z_i}{\partial \theta_i} < 0 \quad (44)
\]
\[
\frac{\partial a_i}{\partial \theta_i} = \frac{\partial a_i}{\partial z_i} \frac{\partial z_i}{\partial \theta_i} = A(0.5 - \tau_i) \frac{\partial z_i}{\partial \theta_i} < 0 \quad (45)
\]
Recall that according to the derivation of (20)
\[
q_i = \left[ (b + B)CE_{i-1} - Bh^* - \frac{1}{2} z_i A(b + B)CE_{i-1} \right] z_i
\]
since we have assumed \(a = 0\). Using formula (34) we can rewrite the above as
\[ q_t = A(b + B)CE_{t-1} \frac{v(1 - \tau_t)}{\gamma} z_t^2 \tag{46} \]

which implies, according to (43), that

\[ \frac{\partial q_t}{\partial \theta_t} < 0 \tag{47} \]

for \( t=t_0 \). Combining (43) and (46) and referring to (10) we can conclude that for \( t=t_0 \)

\[ \frac{\partial E_t}{\partial \theta_t} < 0 \tag{48} \]

We can now proceed to the next step of the induction and evaluate the effect born by the education policy variables in period \( t=t_0+1 \), keeping in mind two sources of this effect: the direct effect of higher value of \( \theta_{t+1} \) and the indirect one caused by lower education quality in the previous period \( E_{t_0} \) as established in (48). The results in (43), (47) and (48) show that the direct effects on the variables \( z_{t+1}, q_{t+1}, E_{t+1} \) in any period \( t+1 \) of a contemporaneous rise in \( \theta_{t+1} \) are negative. For the purposes of completing the induction argument it will therefore be sufficient to prove that a decline in \( E_t \) will have a negative effect on \( z_{t+1}, q_{t+1}, E_{t+1} \), in other words that the derivatives of these variables with respect to \( E_t \) are all positive.

According to (33) the derivative \( \frac{\partial (z_{t+1}^2)}{\partial E_t} \) has the same sign as the expression

\[ 1 - \frac{\theta_{t+1} Bh^*}{(\theta_{t+1} (b + B) - 1) ACE_{t-1}} \]

which according to Lemma 3 is equal to \( 1 - A^{-1}a^* \) and therefore positive. Thus we can conclude that

\[ \frac{\partial z_{t+1}}{\partial E_t} > 0 \tag{49} \]

Rewriting expression (46) for period \( t+1 \): \( q_{t+1} = A(b + B)CE_t \frac{v(1 - \tau_{t+1})}{\gamma} z_{t+1}^2 \), we note that a rise in \( E_t \) affects \( q_{t+1} \) directly (obviously positively) as well as indirectly
through $z_{t+1}$, also positively according (49). We can therefore conclude that $\frac{\partial q_{t+1}}{\partial E_t} > 0$.

This combined with (49) implies due to (10) that $\frac{\partial E_{t+1}}{\partial E_t} > 0$. Thus according to the above discussion the Theorem’s proof is complete. ⁶
Appendix 2. Glossary of Mathematical Terms

\( \beta \)  
Discount factor in individual intertemporal preferences

\( \tau_t \)  
Labor income tax rate in period \( t \)

\( T_t \)  
Total government revenue in period \( t \)

\( a \)  
Lower bound on innate ability, from Section 4 on \( a = 0 \) is assumed

\( A \)  
Upper bound on innate ability

\( a(\omega) \)  
The innate ability of individual \( \omega \)

\( a^*_t \)  
Ability cut-off level for attending college in period \( t \)

\( h^u(\omega) \)  
Unskilled individual’s level of human capital in period \( t \)

\( H^u_t \)  
Aggregate unskilled human capital in goods production in period \( t \)

\( h^s_t(\omega) \)  
Skilled individual’s level of human capital in period \( t \)

\( H^s_t \)  
Aggregate skilled human capital in period \( t \)

\( H^{su}_t \)  
Aggregate skilled human capital in goods production in period \( t \)

\( C \)  
Productivity coefficient of compulsory basic education

\( E_t \)  
Public education quality in period \( t \)

\( b \)  
Coefficient of in-college depreciation of pre-college human capital.

\( B \)  
Productivity coefficient of higher education

\( h^* \)  
Human capital threshold for admission to college

\( D \)  
TFP coefficient in the goods production sector

\( \alpha \)  
Physical capital income share in the goods production sector

\( \theta_t \)  
Productivity augmentation of skilled human capital (skill premium) in the production sector in period \( t \)

\( \gamma \)  
Returns to quantity of teachers

\( \nu \)  
Returns to quality of teachers

\( \Sigma_t \)  
Set of individuals employed as teachers in period \( t \)

\( z_t \)  
Number of teachers (the share of teachers in the working population) in period \( t \)
$I_t^h$  Teacher’s salary in period $t$

$a_t$  The lowest ability level among teachers in period $t$

$h_t$  The lowest human capital level among teachers in period $t$

$\bar{a}_t$  The highest ability level among teachers in period $t$

$\bar{h}_t$  The highest human capital level among teachers in period $t$

$q_t$  Aggregate teacher quality in period $t$
Chapter 2
Self-selection, Unions, and Teacher Attrition

Abstract

Teacher attrition rates in non-unionized schools are two to three times higher than in unionized schools. It would seem that teachers’ unions are the main factor responsible for the difference in attrition rates due to the dismissal barriers provided by teachers’ unions. In a Bayesian framework with Markov Chain Monte Carlo (MCMC) methods, I estimate how much of the propensity to remain in teaching can be attributed to the effect of teachers’ unions rather than the characteristics of individuals who tend to work in unionized schools. With the use of the counterfactual, I find that while teachers’ unions do lower the attrition rate, most of the difference in attrition rates should be attributed to the self-selection into the teaching occupation and not to the barriers to dismissal unions provide.

*JEL* classification: C11; C31; I22; J45

Keywords: Teacher Attrition; Teachers’ Unions; Bayesian Analysis, Causal inference; Markov Chain Monte Carlo; Potential outcomes; Roy Model; Treatment effects
1. Introduction

Teacher attrition rates in unionized and non-unionized schools are distinctly different. After the first year of teaching, 8.3% of teachers in unionized schools exit the teaching force to pursue a career outside of full-time teaching compared to 16.7% of teachers in non-unionized schools (author’s calculations). Teachers’ unions often provide lengthy dismissal procedures with mandatory multiple teacher assessments, teachers’ appeals, and personal development plans in an effort to improve the teacher’s quality prior to dismissal. Non-unionized schools seldom have such constraints and dismiss poor performing teachers when required. Thus, it would seem that teachers’ unions account for the difference in attrition rates due to the dismissal barriers provided by teachers’ unions.

The ways individuals enter into the teaching profession are also distinctly different. The majority enter teaching by attending schools of education for a career as a teacher in a public, unionized, school. Some choose to enter teaching under provisional licensing and emergency certification. Such individuals are under remediation until they have satisfied the necessary course requirements and classroom experience to obtain a regular teaching certificate. These positions are predominantly in needy inner-city public schools where license requirements are demanded by the union. Others choose to enter teaching without any formal teaching coursework or classroom experience, and do not hold any teacher certifications. These individuals work predominantly in non-unionized schools. Given these options, individuals self-select themselves into a certain track.

The self-selection process is determined by the type of school which best fits their individual preferences and abilities. Administrators in unionized schools are bound by collective bargaining contracts which stipulate salary schedules, teacher credential requirements, and tenure. Those who desire more job security, or a teaching position which relies more on teaching credentials than subject knowledge, may find that unionized schools provide more benefits than non-unionized schools. Non-unionized schools provide little permanent job security but do provide teachers with individual specific salaries and more diverse subject matter. Those who have advanced degrees in their subject may prefer to teach in non-unionized schools which unlink their salaries from a schedule. Similarly, those with religious preferences or who tend towards more
non-traditional learning methods may find that non-unionized schools provide a more suitable environment for them than unionized schools.

The self-selection process commences by the choice of college track. Individuals attending college have opportunities to increase scholastic knowledge or receive teaching tools through various majors. An individual’s education path is, for the most part, determined by their scholastic ability. Those with lower scholastic scores tend to choose an education path which makes them eligible to work in union teaching schools where they can earn more under the collective bargaining agreement than under an individual based salary. Those with higher scholastic scores tend to choose a scholastic education path which provides individual based salaries. If they desire to teach, they tend to do so in non-unionized schools where teaching credentials are minimal.

The self-selection into teaching also affects an individual’s duration in the teaching profession. Teachers who work in public unionized schools must satisfy strict entry requirements and, having had previous education courses and classroom experiences, are more likely to be better prepared to teach and understand their job requirements. They also tend to have degrees from schools of education specializing in teaching. Switching professions is costly, as the skills from an education degree are not readily transferable to other professions. Non-unionized school teachers have less to lose from exiting teaching. By not being required to hold a teaching certification or teaching courses, individuals are able to try out teaching while leaving the door open for other careers. Thus, the difference in teacher attrition observed in the data may have to do more with individuals’ self-selection into teaching than to the barriers to dismissal unions provide.

The purpose of this paper is to estimate how much of the propensity to remain in teaching can be attributed to the effect of teachers’ unions rather than to the characteristics of individuals who tend to self-select in unionized schools. I specifically investigate the difference in a teacher’s propensity to exit teaching if she teaches in a unionized school rather than in a non-unionized school. That is, given individuals’ characteristics and preparation are fixed immediately prior to entry into the teaching profession, I estimate how much of their propensity to exit teaching is due to teaching in
a unionized school for all teachers. A potential outcomes model provides the necessary framework to estimate these propensities where Bayesian techniques permit the explicit estimation of the counterfactual attrition rates. The empirical strategy is to estimate each teacher’s propensity to exit teaching in unionized and non-unionized schools. By explicitly calculating each teacher’s propensity to remain teaching in both unionized and non-unionized schools, the effect of unions on attrition is empirically available. If an individual’s propensities are identical, then unions have no effect and the entire propensity to remain in teaching is attributed to individual’s characteristics. If the propensities are different, then the difference is attributed to the union. There are many characteristics of teachers that are unobservable that influence teachers’ entry and exit decisions. This issue is alleviated using a potential outcomes model. This class of models account for the correlated decision of entry into teaching and the exit decision.

Much of the literature on the teacher labor supply decisions has focused on which influences affect unionized public school teacher attrition and, specifically, on the decision to exit teaching, with little regard as to how these individuals entered teaching. Boyd et al. (2008) study the attrition rate of first year teachers based on the teachers’ value-added to student achievement. They find that New York City public school teachers from 2000 to 2005, which they identify as less effective at improving test scores, have higher attrition rates than more effective teachers. Stinebrickner (2002) using NLS-72 data finds that the fertility decision is an important determinant of teacher attrition for female teachers. Hanushek et al. (2003) find that teacher mobility is related more to student characteristics than salary. Ondrich et al. (2005) find teachers with higher salaries relative to local nonteaching salaries are less likely to leave teaching. Similar results are found in Imazeki (2005). In this paper I focus on both unionized and non-unionized schools’ teachers in their first year of teaching when the gap in attrition rates is greatest. The first year is also the only year when all teachers in the U.S. have not received tenure.  

---

8 Mississippi and North Dakota are exceptions as they only require one year of teaching prior to assessment for a continual contract (tenure status). However, in Mississippi, the decision to terminate can be appealed at least twice to the chancery and supreme courts (Mississippi Statute 37-9-59). North Dakota doesn’t have such appeals.
Using teachers’ estimated propensities to exit teaching, I find that 91% of all teachers’ exit decision is independent of union status. That is, 91% exit or remain teaching regardless of whether they teach in a unionized school or not. 6.4% of all teachers remain teaching in unionized schools but exit teaching in non-unionized schools. The remaining 3% are more likely to remain teaching in non-unionized schools but exit teaching when placed in unionized schools. This provides evidence that the difference in teacher attrition observed in the data has more to do with individuals’ self-selection process into teaching than with the barriers to dismissal unions provide. Using the propensities to exit teaching in probabilities, I estimate that 21% of propensity to remain in teaching can be attributed to teachers’ unions while 79% of the propensity to remain in teaching is due to the individual characteristics of teachers. Thus, even though unions affect the majority of teachers’ propensities to exit teaching, teachers’ decisions to exit are not altered enough by unions to explain the difference in attrition rates observed in the data.

The paper is structured as follows: in Section 2 I provide a simple framework to analyze teachers’ entry and exit decisions; Section 3 concerns the data; Section 4 lays out the empirical model; Section 5 provides results; Section 6 concludes. Throughout this paper I will refer to teachers in unionized school as unionized teachers and similarly teachers in non-unionized schools will be referred to as non-unionized teachers. I do this for clarity in exposition.

2. The Determinants of First Year Teachers’ Mobility Decisions
Hanushek et al. (2003) set up a school choice model which provides a framework for analyzing the mobility decisions of established teachers. I extend this framework to include the initial decision of entering teaching. Consider the following stylized two period occupational choice problem. Individual \( i \) is endowed with scholastic aptitude, \( a(i) \), and teaching skills, \( t(i) \). Assume that all individuals under consideration have at least a college eligible level of scholastic aptitude and that individuals know their scholastic aptitude but do not fully know their teaching skills. Rather, individual \( i \) forms
an expectation of teaching skills. Individual $i$'s expectation is $E(t_i)$ with precision $\text{var}(t_i)$. Teaching skills are defined as the productivity level in teaching. Individuals gain knowledge of their teaching skills through classroom experience.

The first decision is the choice of major. There are two types, one which enhances scholastic aptitude and another to improve teaching skills. Individuals must choose only one type. The per unit cost is identical. If a scholastic track is chosen, aptitude is augmented to $a'(i) \geq a(i)$. If a teaching path is chosen, classroom experience (a practicum) augments teaching skills to $t'(i) \geq t(i)$, and are qualified to teach in a unionized school. Due to the practicum, those who choose a teaching path during education are more informed of their teaching skills, $\text{var}(t_i) \leq \text{var}(t_i)$.

The second decision is the choice of career. Individuals choose one of three career paths: (1) private sector worker, (2) non-union teacher, or (3) union teacher. All individuals are eligible to work in the private sector and earn a wage based on their scholastic aptitude and working experience. A new private sector worker earns $w_1(a(i))$ and will earn $w_2(a(i)) > w_1(a(i))$ with one period of experience. This reflects the increase in wage from gaining experience. Individuals with sufficient scholastic aptitudes, $a_{nu}$, can work as teachers in non-unionized schools. New non-unionized teachers earn a wage $w_1(a_{nu}) < w_1(a(i))$ and $w_2(a(i), t(i))$ with one period experience. Upon completion of non-unionized teachers’ first year, the non-unionized school administrators gain knowledge of their teaching skills in transmitting their scholastic ability. The second period wages can potentially be higher than second period private sector wages. New unionized teachers earn a wage of $\tilde{w}_1 > w_1(a_{nu})$ as a new teacher and $\tilde{w}_2$ in the second period. The higher wage reflects the pre-screening taken place to ensure some teaching quality and that an additional skill component adds value to these workers. Unionized teachers are screened prior to placement and only those individuals with sufficient teaching skills, $t_u$, can work as unionized school teachers.

After education is obtained, individual $i$ chooses an occupation, $d^*$, from among the feasible opportunities in order to maximize the present value of expected utility:
\[
\max_{d'} PV[u(X_d(i), Z(i), L)] \\
given d' \in \{d\}
\]

where \(\{d\}\) is the set of eligible occupations, \(X_d(i)\) are the job characteristics including working conditions and wages of job \(d\), \(Z(i)\) captures individual factors, and \(L\) are government policies regarding teachers’ unions. At the end of the first period, all teachers update their knowledge of teaching skills based on their position. At the beginning of the second period, individuals have an opportunity to choose whether to remain at their current position or switch.

I first analyze the education and occupation decisions and then discuss the mobility decision. For clarity, I analyze the problem by aptitude level and expectation of teaching skills. First, those of low aptitude and expected low teaching skills work in the private sector as they are ineligible for teaching in either a unionized or non-unionized school. These individuals choose a scholastic non-teaching path for schooling. Second, those who have pre-screened teaching skills above \(u_t\), but low aptitude, qualify for unionized school teaching or private sector employment. All choose a teaching path for schooling and unionized school teaching for employment, as this will offer them higher salary than private sector employment. Third, as aptitude rises, individuals have two or three occupational choices depending on expected teaching skills. These individuals’ education/occupational decisions are more complex. Those with low teaching skills and who are quite precise in their expectations forgo teaching in non-unionized schools and teach in a unionized school. This is because second period non-unionized teaching salaries are determined by scholastic ability and teaching skills. Individuals choose a teaching path for education. If individuals’ expectation of teaching skills has a low precision, then they teach in a non-unionized school in the first period and choose a scholastic path instead of a teaching path for education to insure against incorrect expectations. If their realization of teaching skills after the first period is lower than expected, they can earn more in the private sector in the second period, having augmented their scholastic ability in school. Lastly, those with the highest ability choose between non-unionized teaching and private sector employment. All of these individuals
choose a scholastic path for education and only work as non-unionized teaching if their expectations on their teaching skills are high.

Now consider individuals who have completed their first year of teaching and have updated their beliefs on the characteristics of teaching jobs. Individuals may also have had personal changes in their lives which can affect their job switching. Let individual $i$ have job switching costs from current job to a new job be $c_d$ for job opportunity $d$. The individual will switch jobs if the expected utility gain is sufficient to overcome the cost.

There are several costs to switching jobs. These costs depend on the education path chosen and on whether the individuals were able to update their expectations on their teaching skills. That is, their expectations on teaching skills are dependent on the teaching experience during that year. These updates are not uniform across all teachers, as each school and classroom provides different experiences. Once these expectations are updated, an individual can calculate the costs of job switching. I analyze the switching costs by educational path and occupation chosen during the first period. If an individual chooses a scholastic path and a private sector career, then switching costs are similar to that of the first period but with a higher private sector wage. This is because private sector workers’ expectations of teaching skills remain the same, i.e., they are not updated. If an individual chooses a scholastic path and a non-unionized teaching position, then switching costs are dependent on the individual’s updated expectations on teaching jobs. If expectations change substantially then, depending on which way they change, they either reinforce the initial occupational decision or provide disincentives to remain in teaching. When updated expectations are such that teaching is worse than expected, individuals have low or even negative switching costs. If an individual chooses unionized teaching, then the switching costs are greater than non-unionized teachers. This is due to two reasons: 1) their scholastic aptitude was not augmented during schooling and as such, if they work in the private sector, they will earn lower salaries; 2) unionized school teaching salaries are not directly linked to scholastic ability or teaching skills while switching to a new occupation would link their pay to scholastic ability.
There are other costs an individual considers when deciding to switch careers. Some include changes in school personnel policies and working conditions, $X_d$, changes in the individual’s characteristics, $Z(i)$, and changes in government policies regarding teachers’ unions, $L$. I discuss these in more depth below.

School administrators may change an individual’s teaching load, the courses to be taught, or even the location where they will be teaching the following year. These may have profound influences on an individual’s decision to remain in teaching. Measuring teachers’ salaries also poses a problem. Depending on the length of year and extracurricular activity involvement, salaries may be quite different; without a standard measure across teachers. To address this, I assume individuals choose employment based on the total salary. That is, they choose between jobs which package a total salary for certain duties. This seems quite intuitive as many prospective new teachers can’t \textit{a la carte} their working conditions. Working conditions also include standard measures such as class size and type of teaching assignment. To address general working conditions, I also control for the type of school, religious affiliation held, and the level of the school the teacher works in. Community characteristics may also play an important role. Those set in the rural versus urban may find it difficult to retain teachers, as teachers may desire to work in more affluent communities and be less likely to leave upon receipt of such an assignment. Furthermore, educated communities may provide more support as well as better working conditions for teachers.

Teacher’s preferences also play a role in a teacher’s decision to switch jobs. Marriage and children may influence individuals’ decisions. This seems quite plausible since many teachers exit the work force to care for family, as discussed above. Teachers’ safety may also play an important role. Some individuals may have aspirations to change the world and choose a school where they feel they can do the most good. These aspirations may be outweighed by the reality of their personal safety once they are attacked or threatened that they will be attacked. Type of education may also influence an individuals’ mobility decision. Those with a strict Bachelor’s degree in education may find that their outside options are limited, as the private sector doesn’t value teachings skills as much as the skills other degrees offer.
State legislation regarding teachers’ unions can affect an individual’s decision to join a teachers’ union. If a state does not require a union, then individuals are more inclined to choose a scholastic path since, if their expectations on teaching skills are incorrect, they can earn a higher salary in the second period. I use state legislation regarding collective bargaining laws to identify individuals’ decisions to enter a unionized school versus entering a non-unionized school.

3. The Data
The main data source is the restricted-access Teacher Follow-up Survey (TFS) years 2000-01 and 2004-05. Each TFS consists of two questionnaires, one for current teachers and one for former teachers. The TFS consists of a sample of teachers from the Schools and Staffing Survey (SASS) from the previous year. The TFS primary design is to measure attrition and provide comparisons on teachers on many different dimensions who left the teaching profession, teachers who moved to another school, and those who stayed in the same school. It also contains current economic activities of recently exited teachers. I also use community characteristic data from the U.S. Census and other ancillary datasets. Appendix A describes the dataset construction, the variable link between the TFS datasets, and also contains the restrictions placed on variables.

The final data set consists of 1080 full-time regular first year teachers. There are 467 unionized teachers and 613 non-unionized teachers. Applying the sampling weights, the weighted observations are 215,760 unionized teachers and 99,368 non-unionized teachers. Included in the non-unionized teachers are a significant portion of public school teachers who identify themselves as not being a member of a teachers’ union. Table A3 in Appendix A contains the full descriptive statistics for all variables used in the analysis.

The sample consists of 75% females with the average age being 29 years old. 82% of the sample is non-hispanic white, and a little less than half are currently married. The average number of children per teacher who are under the age of 5 is .33, and half the teachers live in households with incomes above $50,000. Few teachers have outside jobs during the school year, with this being less so for non-unionized school teachers.
The average class size is 15 in non-unionized schools, with one additional student per class in unionized schools. 10% of all first year teachers reported they were either threatened or attacked over the previous year, with more unionized teachers being threatened/attacked than non-unionized teachers. 91% of unionized teachers had a teacher practicum, compared with 76% of non-unionized teachers. Non-unionized teachers have higher ACT scores and are more likely to teach math or science. Roughly half of all first year teachers were previously unemployed prior to accepting a position in teaching. Charter schools are 5 times more likely to be non-unionized. Lastly, 77% of all unionized teachers have undergraduate degrees in states which require collective bargaining contracts, compared to 56% of non-unionized teachers.

There are two dependent variables in the analysis, one for the treatment equation: being a member of a teacher’s union; and one for the outcomes equations: the exiting decision of teachers. They are both limited-dependent variables. Overall, the weighted sample consists of 68.5% unionized school teachers. The mobility decisions of teachers are displayed in Table 1. Most teachers remain as full-time teachers at the same school where they taught the previous school year. Those who switch to a new school, and continue on as full-time teachers, did so at the same rate regardless of union membership. The difference in mobility between unionized and non-unionized teachers is apparent in those who left teaching. Non-unionized teachers exit the profession at twice the rate of unionized teachers.

<table>
<thead>
<tr>
<th></th>
<th>Union</th>
<th>Non-Union</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teaching at the Same School</td>
<td>78.6</td>
<td>69.0</td>
</tr>
<tr>
<td>Teaching at a New School</td>
<td>13.1</td>
<td>14.3</td>
</tr>
<tr>
<td>No Longer Teaching</td>
<td>8.3</td>
<td>16.7</td>
</tr>
</tbody>
</table>

Source: TFS 2004-05 (STTUS_TF, TFSFINWT) and TFS 2000-01 (STATUS, TFSFINWT)

Footnotes
1: In percent.
2: Weighted observations

This is robust across the teacher experience profile.
Some of the difference in teacher attrition may be attributed to the difference in barriers to dismissal. Table 2 provides a snapshot of different barriers and to what extent they are problematic for administrators at unionized and non-unionized schools. The barriers to dismissal are roughly three to four times higher for unionized schools than non-unionized. For example, 76% of all unionized school principals indicated that tenure was a barrier to dismiss inadequate or poor teachers compared to 16% of non-unionized school principals.

<table>
<thead>
<tr>
<th></th>
<th>Union</th>
<th>Non-union</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tenure</td>
<td>76</td>
<td>16</td>
</tr>
<tr>
<td>Teachers’ Associations</td>
<td>69</td>
<td>0</td>
</tr>
<tr>
<td>Inadequate Documentation</td>
<td>51</td>
<td>27</td>
</tr>
<tr>
<td>Personnel policies</td>
<td>52</td>
<td>19</td>
</tr>
<tr>
<td>Dismissal is too stressful and uncomfortable</td>
<td>25</td>
<td>16</td>
</tr>
<tr>
<td>Termination decisions not upheld</td>
<td>24</td>
<td>7</td>
</tr>
</tbody>
</table>

Table 2: Schools’ Barriers to Dismissal of Teachers<sup>1,2</sup>

Source: SASS 2003-04 (A0152-A0159, D0094, AFNLWGT)
SASS 1999-00 (A0174-A0179, D0497, S0497, AFNLWGT)

Footnotes
1: In Percent
2: Weighted Principal level observations

The degree of barriers to dismissal also vary from state to state for public unionized schools.<sup>10</sup> Most states require two evaluations prior to commencing the dismissal process. These may be six months to one year apart. Teachers receiving poor evaluations can appeal the actual evaluation or receive leniency given through collective bargaining agreements. For instance, in North Carolina, teachers on remediation are placed on a year-long summative evaluation process with remediation (North Carolina Department of Public Instruction, TPAI Fact and Summary Sheet). In Georgia, the decision to terminate a teacher may be appealed at least twice to the state board and to the superior court of the county (Georgia Code 20-2-1160). In Massachusetts, if a provisional teacher's performance is deemed insufficient, the teacher may enter a district training program prior to receiving tenure (Massachusetts General Law Title XII, ch. 71, sec. 38).

<sup>10</sup>The TR3 database constructed by the National Center on Teacher Quality contains state level legislation on barriers to dismissal.
Similarly, Louisiana exempts provisional teachers from evaluations while participating in the *Louisiana Teacher and Assessment Program* (State BP title 28, Part CIII, Bulletin 1525, Ch 3, #345). Furthermore, some teacher terminations are not upheld by senior administrators or administrators may extend a teachers’ contract if their evaluations are unacceptable. In Ohio, the superintendent can make a one-time only recommendation for extension of the probationary term up to 2 years (Ohio Revised Code 3319.11 (C)(1)).

*Table 3* displays the most important reasons teachers indicated in deciding to exit after their first year of teaching. 27.9% of unionized teachers versus 30.4% of non-unionized former teachers left due to outside opportunities while 22.7% of unionized versus 19% of non-unionized former teachers left due to school staffing actions. 19% of unionized versus 26% of non-unionized former teachers left due to personal reasons.

<table>
<thead>
<tr>
<th>Table 3: Most Important Reason Why Teacher Says They Leave</th>
<th>Union</th>
<th>Non-Union</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teaching Job Expectations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dissatisfied with Teaching as a Career</td>
<td>15.4</td>
<td>11.4</td>
</tr>
<tr>
<td>Dissatisfied with Teaching Assignment</td>
<td>14.4</td>
<td>4.6</td>
</tr>
<tr>
<td>Outside Opportunities</td>
<td></td>
<td></td>
</tr>
<tr>
<td>For Better Salary</td>
<td>23.1</td>
<td>19.8</td>
</tr>
<tr>
<td>To Pursue position outside of K-12 Teaching</td>
<td>4.8</td>
<td>10.6</td>
</tr>
<tr>
<td>Personnel Policies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>School Staffing Action</td>
<td>22.7</td>
<td>19.0</td>
</tr>
<tr>
<td>Personal Reasons</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Child Rearing</td>
<td>---</td>
<td>12.8</td>
</tr>
<tr>
<td>Schooling (in education)</td>
<td>8.7</td>
<td>1.4</td>
</tr>
<tr>
<td>Schooling (outside of education)</td>
<td>---</td>
<td>8.8</td>
</tr>
<tr>
<td>Health</td>
<td>8.1</td>
<td>0.4</td>
</tr>
<tr>
<td>Personal Reasons</td>
<td>2.2</td>
<td>2.7</td>
</tr>
</tbody>
</table>

Source: TFS 2004-05 (f0579)

Footnotes
1: Weighted observations.

While *Table 3* indicates why teachers left, it does not provide any substantial proof that unions have any impact on teachers’ decisions to exit teaching. To more fully understand the outside opportunities for teachers, I analyze the careers chosen by former teachers and present the results in *Table 4*. The occupational paths chosen by former teachers are quite different between former unionized and non-unionized teachers. 43% of former unionized teachers remain in the education field with almost all of them working as general staff. This is compared to 20.5% of non-unionized former teachers.
who remain in the education field. The proportion of former teachers who work outside of the education field is quite similar between unionized and non-unionized teachers, 30% of former unionized teachers versus 36% of former non-unionized teachers and these working predominantly as office and administrative support. 20% of former unionized teachers choose to exit the labor force compared to 23% of non-unionized. Interestingly, 8% of unionized and 21% of non-unionized classified themselves as unemployed.

| Table 4: Occupational Status of Former Teachers With 1 Year Experience\(^1,2\) |
|-------------------------------------|-------|-------|
| Occupations Inside of Education     | Union | Non-Union |
| General Staff\(^3\)                 | 30.7  | 20.1  |
| Administrator\(^4\)                | 0.0   | 0.2   |
| Guidance Counselor                  | 12.2  | 0.2   |
| Occupations Outside of Education\(^5\) | 30.2  | 36.1  |
| Office and administrative support occupations | 10.2  | 5.8   |
| Computer and mathematical science occupations | 3.9   | 4.5   |
| Business and financial operations occupations | 2.8   | 5.7   |
| Sales and related occupations       | 4.9   | 3.4   |
| Management occupations              | 0.0   | 4.1   |
| Arts, design, entertainment, sports, and media occupations | 1.8   | 2.0   |
| Community and social service occupations | 2.9   | 0.5   |
| Building and grounds cleaning and maintenance occupations | 0.0   | 2.4   |
| Healthcare practitioner and technical occupations | 0.0   | 2.2   |
| Healthcare support occupations      | 1.0   | 0.8   |
| Not in the Labor Force              | 20.0  | 22.9  |
| Caring for Family Member            | 0.7   | 8.1   |
| Student                             | 18.7  | 14.7  |
| Retired                             | 0.2   | 0.0   |
| Disabled                            | 0.4   | 0.0   |
| Unemployed                          | 7.3   | 21.0  |

Source: TFS 2004-05 (OCODE_TF F0553 F5556 TFSFINWT), TFS 2000-01 (F0053 F5055 TFSFINWT)

Footnotes
1: In percent.
2: Weighted observations.
3: General staff include teachers’ assistants, secretaries, bus drivers, janitorial staff, and other non-teaching positions.
4: Administrators include deans, principals, assistant and vice principals, and supervisory staff.
5: Occupations are classified by 2002 NAICS Occupation Codes.

Table 5 contains the average earnings of former first year teachers when they were teachers and then for their new occupations. Former union teachers make $5,623 more than former non-union teachers in the teaching profession. In their new
occupations, earnings are roughly the same between union and non-union teachers at $32,000. Thus, former union teachers’ earnings remain roughly the same, while former non-union teachers’ earnings rise substantially.

<table>
<thead>
<tr>
<th>Table 5: Average Earnings of Former First Year Teachers¹,²,³</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Previous Year’s Earnings (As a Teacher)</td>
</tr>
<tr>
<td>Current Year’s Earnings (New Occupation)</td>
</tr>
<tr>
<td>Difference In Earnings</td>
</tr>
<tr>
<td>Lowest Salary to Return to Teaching</td>
</tr>
<tr>
<td>Difference In Earnings As a Teacher</td>
</tr>
</tbody>
</table>

Source: TFS 2004-05 (earnsch, F059, F0612), TFS 2000-01 (earnsch, F0058), and BEA GDP Deflator

Footnotes
1: In 2004 dollars.
2: Excludes unemployed or not in labor force.
3: Weighted observations.

Former non-union teachers have a higher reservation wage to return to teaching, as indicated in their lowest salary to return to teaching. That is, non-unionized teachers require a raise of $8,653 to enter back into teaching, compared to former unionized teachers requiring a salary raise of $946 to enter back into teaching. Even though these individuals may overvalue their teaching skills to what the market would bear, the non-union teachers’ reservation wage being over $4000 above that which they make in the private sector indicates that they prefer private sector work.

Comparing teachers who remain in teaching with those who exit, both unionized and non-unionized teachers who remain in teaching make roughly $2,500 more than those who exit. This indicates that earnings may partially explain why both union and non-union teachers leave. If the pay distribution across non-union schools is more disburse than non-unionized pay, then it is possible that the higher attrition rate is due solely to the pay disparity. Union schools are less likely to have a high pay disparity since unions bargain for collective wages across larger pools of teachers. After the first year, salaries between union and non-union teachers converge with non-unions teachers gaining $648 while the union teachers lose $426. Signing bonuses and other perks to entice highly demanded teachers into union schools may explain why they drop in salary. This pattern continues over the experience profile.
Table 6: Average Earnings of First Year Teachers Who Remain Teaching In Their Second Year$^{1,2}$

<table>
<thead>
<tr>
<th></th>
<th>Union</th>
<th>Non-Union</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>First Year Earnings</td>
<td>35,602</td>
<td>29,940</td>
<td>5,662</td>
</tr>
<tr>
<td>Second Year Earnings</td>
<td>35,176</td>
<td>30,588</td>
<td>4,588</td>
</tr>
<tr>
<td>Difference in Earnings</td>
<td>-426</td>
<td>648</td>
<td></td>
</tr>
</tbody>
</table>

Source: TFS 2004-05 (earnsch, earnsh_tf), TFS 2000-01 (earnsch, earnsc_tf), and BEA GDP Deflator
Footnotes
1: In 2004 dollars.
2: Weighted observations.

When analyzing first year teachers’ pay, it is important to recognize that unions pay more due to an initial screening of teaching skills. Union teachers have a higher proportion of teachers who hold teacher certification. This screens poor teachers from entering into teaching, only to be dismissed after the first year.

Table 7: Certification Type by Union Status and Mobility Decision of First Year Teachers$^1$

<table>
<thead>
<tr>
<th></th>
<th>Union</th>
<th>Non-Union</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Remain in</td>
<td>Exit</td>
</tr>
<tr>
<td></td>
<td>Teaching</td>
<td>Teaching</td>
</tr>
<tr>
<td>Regular License</td>
<td>59.5</td>
<td>46.2</td>
</tr>
<tr>
<td>Probationary Certificate</td>
<td>17.9</td>
<td>18.9</td>
</tr>
<tr>
<td>Provisional License</td>
<td>12.5</td>
<td>26.0</td>
</tr>
<tr>
<td>Temporary, Waiver, or No Certification</td>
<td>10.1</td>
<td>9.0</td>
</tr>
</tbody>
</table>

Source: TFS 2004-05 (T0443, T0166, TFSFINWT) and TFS 2000-01 (T0104, TFSFINWT)
Footnotes
1: Weighted observations

Table 7 contains certification types of first year teachers by union status and mobility decision. The data suggests that requiring a teacher practicum for union teachers lowers the attrition rate, since it serves both the prospective teacher and potential employers. In-class experience aids the prospective teacher in understanding if teaching is a desired career. Potential employers are provided with a screening mechanism for teacher quality. For unionized teachers, 59.5% of those teachers who remained had a regular license compared to 46.2% who exited. Non-union teachers who remained or left were equally likely to hold a regular license. This suggests that obtaining a teaching license in a non-union position isn’t a factor of attrition.
Teachers’ prospects of returning to teaching are found in Table 8. 30% of unionized and non-unionized former teachers indicate that they would return the following year if they returned at all. However, 42.1% of former unionized teachers are undecided compared to 23.3% of former non-unionized teachers. It should be noted that the category never was excluded from the choices provided to teachers surveyed. Thus, the undecided category could indicate that former teachers are indicating they don’t have a set date of returning or they will never return to teaching.

<table>
<thead>
<tr>
<th></th>
<th>Union</th>
<th>Non-Union</th>
</tr>
</thead>
<tbody>
<tr>
<td>Later This School Year</td>
<td>13.2</td>
<td>6.3</td>
</tr>
<tr>
<td>Next School Year</td>
<td>29.0</td>
<td>30.6</td>
</tr>
<tr>
<td>In Less Than Five Years</td>
<td>13.2</td>
<td>30.7</td>
</tr>
<tr>
<td>After Five Years</td>
<td>2.6</td>
<td>9.1</td>
</tr>
<tr>
<td>Undecided</td>
<td>42.1</td>
<td>23.3</td>
</tr>
</tbody>
</table>

Table 8: Expected Return Date of Former Teachers

Source: TFS 2004-05 (f0613) and TFS 2000-01 (f0613)

Footnotes
1: Weighted observations.

4. The Model

The modified potential outcomes model is an extension of the Roy (1951) model. This framework allows the explicit modeling of the counterfactual state – the outcome that would have been observed had the individual made a different treatment decision. This class of models has seen considerable use in the program evaluation literature, where the goal is to determine the effectiveness of a program or treatment when individuals are not randomly assigned into the treated or untreated states. I develop a parametric model to allow endogenous selectivity to investigate the impact of teaching in a unionized school, denoted $D$ on the observed mobility outcome, denoted $y=0$ for staying at current school

---

11 In recent years, Bayesian techniques in this literature have grown immensely with the works of Vijverberg (1993), Koop and Poirier (1997), Li (1998), Chib and Hamilton (2000, 2002), Poirier and Tobias (2003) and Li, Poirier and Tobias (2004).
and y = 1 for exiting teaching. Let \( y_1 \) denote the outcome received by the individual in the treatment state and \( y_0 \) denote the outcome received without treatment. The variables \( y_1 \) and \( y_0 \) are *outcomes variables* and represent either observed or potential outcomes. When \( d = 1 \), the teacher works in a unionized school and thus her ‘treated’ mobility outcomes \( y_1 \) is observed. Conversely, when \( d = 0 \), the teacher works in a nonunionized school, and thus her untreated mobility outcome \( y_0 \) is observed. The observable conditions for teacher \( i \) are

\[
\begin{align*}
y_{i1} &= \begin{cases} y_i & \text{if } d_i = 1 \\ -y_i & \text{if } d_i = 0 \end{cases}, \\
y_{i0} &= \begin{cases} -y_i & \text{if } d_i = 1 \\ y_i & \text{if } d_i = 0 \end{cases}.
\end{align*}
\]

Since only one outcome is ever observed for any individual, these conditions summarize the observed outcomes for teacher \( i \) by the following

\[
y_i = d_i y_{i1} + (1 - d_i) y_{i0}.
\]

There are \( n \) observations with \( n_1 \) of them in the treated group and \( n_0 = n - n_1 \) of them in the non-treated group. Reordering and stacking observations, (1) can be written as

\[
\begin{bmatrix}
Y_i \\
Y_{\text{miss}}
\end{bmatrix} =
\begin{bmatrix}
Y_{1n1} \\
Y_{\text{miss}}
\end{bmatrix}.
\]

\[
\begin{bmatrix}
Y_i \\
Y_{\text{miss}}
\end{bmatrix} =
\begin{bmatrix}
Y_{0n0} \\
Y_{\text{miss}}
\end{bmatrix},
\]

\( n \times 1 \) vectors where \( Y_{1n1} \) and \( Y_{0n0} \) are observed outcomes and \( Y_{\text{miss}}^{\text{1}} \) and \( Y_{\text{miss}}^{\text{0}} \) are potential or missing outcomes. The observed treatment decisions \( d \) and the observed and potential outcomes \( y_1 \) and \( y_0 \) are generated by an underlying latent variable representation of the model. The binary treatment indicator, \( d_i \), is related to the latent \( d_i^* \) as follows:

\[
d_i = I(d_i^* > 0) = I(e_{id} > -W_i\gamma)
\]

with \( I(\cdot) \) denoting the standard indicator function equal to one if the event is true and to zero otherwise. Stacking observations, the treatment variable in (4) can be written as

\[12\] There is no assumed ordering among the alternatives. See Li and Tobias (2008) for ordered potential outcomes model specification.
\[ D = \begin{bmatrix} 1_{n_1} \\ 0_{n_0} \end{bmatrix}, \quad (5) \]

a \( n \times 1 \) vector. Similarly, the treated latent variable can be written as

\[ D^* = \begin{bmatrix} d^*_{n_1} \\ d^*_{n_0} \end{bmatrix}, \quad (6) \]

a \( n \times 1 \) vector. Similarly, I introduce independent unobservable latent variables \( y^*_1 \) and \( y^*_{i0} \) related to the outcome variables \( y_{i1} \) and \( y_{i0} \) as follows:

\[ y_{ij} = I(y^*_j > 0) = I[\varepsilon_{ij} > -X_{ij} \beta_j] \quad \text{for } j = 1, 0 \quad (7) \]

with \( I(\cdot) \) again denoting the indicator function. Reordering the observations, the latent variables in (7) can be stacked as

\[ Y^*_1 = \begin{bmatrix} y^*_1_{i1} \\ y^*_1_{i0} \end{bmatrix}, \quad Y^*_0 = \begin{bmatrix} y^*_{0n} \\ y^*_{0n_0} \end{bmatrix}, \quad (8) \]

with similar dimensions and interpretation as \( Y_i \) and \( Y_0 \).

Figure 1 displays an individual’s propensity to remain or exit measured in latent outcome for individual \( i \), \( y^*_i \). When \( y_i = 0 \), this corresponds to a latent propensity \( y^*_i < 0 \) and when \( y_i = 1 \), this corresponds to a latent propensity \( y^*_i > 0 \). The point representing indifference in the decision to remaining or exiting teaching is \( y^*_i = 0 \). As an individual’s propensity to exit increases (decreases), the corresponding value of \( y^*_i \) increases (decreases). Similarly, as an individual’s propensity to remain increases (decreases), the corresponding value of \( y^*_i \) decreases (increases). The probability of exiting teaching is the area under the curve to the right of being indifferent. Thus, it is expected that the distribution of exiting teaching in a unionized schools is to the left of the non-unionized schools for the majority of individuals.
The Type 5 Tobit (see Amemiya (1984)) model is expressed in the five-equation system:

\[ D^* = W \gamma + \epsilon_D \]  
\[ Y_i^* = X \beta_j + \epsilon_1 \]  
\[ Y_0^* = X \beta_0 + \epsilon_0 \]

with equations (4) and (7) completing the system. Equations (10) and (11) are the mobility outcome equations in the treated and untreated states, respectively. The matrix \( W \) is \( n \times k^D \), \( \gamma \) is \( k^D \times 1 \), \( X \) is \( n \times k \), \( \beta_j \) is \( k \times 1 \) (\( j = 1, 0 \)), and the error terms are of the dimensions of the dependent variables as specified above.\(^{13}\) I assume the availability of some exclusion restriction or instrument, i.e., some covariates which enters \( W \) that is not contained in \( X \).\(^{14}\)

I make the following joint normality assumption of the error terms (writing only the lower triangular portion of \( \Sigma \) for notational simplicity):

\[^{13}\]I assume that the same set of covariates appear in the treated and untreated states. This assumption is not necessary but reasonable for this estimation process.

\(^{14}\)Li and Tobias (2008) point out that if the observed \( D_i \) is an element of \( x_i \), then this would be of the form of a standard treatment or causal effect model that only works with observed rather than potential outcomes (See Maddala (1983) for further elaboration into this class of models).
\[
\begin{bmatrix}
\varepsilon_{id} \\
\varepsilon_{i1} \\
\varepsilon_{i0}
\end{bmatrix}
| X_i, W_i \sim \text{N} \left( \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \sigma_{D1} & 1 \\ \sigma_{D1} & \Sigma & \sigma_D \\ 1 & \sigma_D & 4 \end{bmatrix} \right)
\]

(12)

or

\[
\varepsilon_i | X_i, W_i \sim \text{N}(0, \Sigma).
\]

(13)

where \( \Sigma \) is a positive definite matrix. The variance parameters have been normalized to unity for identification purposes since the scale of the latent variables are indeterminate. In the following, I impose this condition through the prior.

The primary problem when estimating the causal effect is one of unobserved confounding, whereby unobservable characteristics of the individual not captured in the model are correlated with both the treatment decision (denoted \( D \)) and the outcome of interest (denoted \( y \)) due to some unmodeled subject-specific factors. The presence of the correlation/covariance terms \( \sigma_{D1} \) and \( \sigma_{D0} \) are added to account for the potential of confounding on unobservables.

4.1. The Likelihood

Given the assumed conditional independence across observations, I can write the likelihood function for this model as:

\[
L(\Gamma; y, D) = \prod_{\{i,D_i=1\}} \Pr(y_{i1} = y_i, d_i^+ > 0 | \Gamma) \prod_{\{i,D_i=0\}} \Pr(y_{i0} = y_i, d_i^* \leq 0 | \Gamma)
\]

\[
= \prod_{\{i,D_i=1\}} \int_0^{\infty} \Pr(y_{i1} = y_i, d_i^+) dd_i^+ \prod_{\{i,D_i=0\}} \int_{-\infty}^{0} \Pr(y_{i0} = y_i, d_i^*) dd_i^*
\]

\[
= \prod_{\{i,D_i=1\}} \int_0^{\infty} \Pr(d_i^+ | y_{i1} = y_i) \Pr(y_{i1} = y_i) dd_i^+ \prod_{\{i,D_i=0\}} \int_{-\infty}^{0} \Pr(d_i^* | y_{i0} = y_i) \Pr(y_{i0} = y_i) dd_i^*
\]

where \( \Gamma = [\beta, \Sigma] \). The conditional and marginal densities in these expressions can be determined from the assumed normality of the error vector. For example,

\[
\Pr(y_{i1} = y_i, d_i = 1 | \Gamma) = \Pr(y_{i1}^+ > 0, \mu_i > -w_i \gamma | w_i, x_i, \Gamma)
\]

\[
= \Pr(\varepsilon_{i1} > -x_i \beta, \mu_i > -w_i \gamma | w_i, x_i, \Gamma)
\]

Performing the required calculations, I obtain
\[ L(\beta, \Sigma) = \prod_{(i,D_i=1)} \Phi \left( \frac{\bar{t}_{iD_i} + \rho_{D_1 D_0} t_{iD_0}}{(1-\rho_{D_0}^2)^{1/2}} \right) \phi(t_{iD_0}) \]
\[ \times \prod_{(i:D_i=0)} \Phi \left[ 1 - \left( \frac{\bar{t}_{iD_0} + \rho_{D_0} t_{i0}}{(1-\rho_{D_0}^2)^{1/2}} \right) \right] \phi(t_{i0}), \]

where
\[ \bar{t}_{iD} = W \gamma \]
\[ t_{i1} = y_i^* - X_i \beta_1 \]
\[ t_{i0} = y_i^* - X_i \beta_0 \]

and \( \phi(\cdot) \) denotes the standard normal density.

Note that the correlation parameter \( \sigma_{10} \) does not enter this likelihood function, and thus it is not identified. This parameter describes the correlation between outcomes across regimes. The reason is that the pair of outcomes \((y_1, y_0)\) are never observed for any individual, and thus, the correlation between these two outcomes will not enter the density function for the observed data (this is quite apparent from eq. (3)).

As shown in Vijverberg (1993), Koop and Poirier (1997), Poirier and Tobias (2003), and Li, Poirier, and Tobias (2003), the positive definiteness of the 3x3 covariance matrix serves to place bounds on \( \sigma_{10} \). These bounds imply that, conditionally,
\[ \sigma_{10} \leq \sigma_{10} \leq \bar{\sigma}_{10}, \]

where
\[ \sigma_{\sigma_{10}} = \sigma_{D1} \sigma_{D0} - [(1-\sigma_{D1}^2)(1-\sigma_{D0}^2)] \]
and
\[ \bar{\sigma}_{\sigma_{10}} = \sigma_{D1} \sigma_{D0} - [(1-\sigma_{D1}^2)(1-\sigma_{D0}^2)]. \]

It should be apparent that these are only functions of identified correlation parameters. Information learned on these parameters spills over and serves to update beliefs about the non-identified correlation parameter via the enforced positive definiteness of \( \Sigma \).
4.2. Bayesian Estimation

Instead of direct evaluation of this posterior, modern Bayesian empirical work makes use of recent advances in simulation methods to carry out a posterior analysis. One simulation device in particular is the Gibbs sampler. This algorithm solves the problem of calculating the posterior moments, quantiles, marginal densities and other quantities of interest by first obtaining a set of draws from the posterior \( p(\Gamma \mid y, D) \). The posterior is obtained by generating a sequence of draws that converge to this distribution. Once convergence has been achieved, the subsequent set of simulated parameters values can be used to calculate the desired quantities. In the Gibbs sampler, a Markov Chain whose limiting distribution is \( p(\Gamma \mid y, D) \) is produced by iteratively sampling from the complete posterior conditionals of the model. In this model, I use well-known diffused priors that can easily be sampled from.

I use data augmentation (Tanner and Wong (1987) and Albert and Chib (1993)) in conjunction with the Gibbs sampler. When data augmentation is used, the posterior is first expanded to include not only the parameter vector \( \Gamma \), but also the latent data \( D^*, y_1^*, \) and \( y_0^* \). Data augmentation in conjunction with the Gibbs sampler simplifies the required posterior calculation since conditioned on the latent data, inference regarding the regression parameters proceeds similarly one with a continuous dependent variable. Given the regression parameters, it is straightforward to obtain draws from the posterior conditional for the latent data.

For the model, this augmented posterior is of the form:

\[
\begin{align*}
  p(D^*, y_1^*, y_0^*, \Gamma \mid y, D) & \propto p(y, D, D^*, y_1^*, y_0^*, \Gamma) \\
  &= p(y, D \mid D^*, y_1^*, y_0^*, \Gamma) p(D^*, y_1^*, y_0^* \mid \Gamma) p(\Gamma)
\end{align*}
\]

with \( p(\Gamma) \) denoting the prior for the parameters of the model. The first term, conditioned on the latent variables and model parameters, the observed responses \( D \) and \( y \) are known with certainty and thus the joint (conditional) distribution of \( D \) and \( y \) is degenerate. The middle term expresses the known trivariate normal density given the joint normality assumption (12). Using the assumed conditional independence across observations, I can rewrite the augmented posterior as follows:
\begin{equation}
p(D^*, y_i^*, y_0^*, \Gamma \mid D, y) \propto p(\Gamma)\prod_{i=1}^{n} \phi_3(s_i; \beta, \Sigma) \times \\
\begin{bmatrix}
I(d_i = 1)I(d_i^* > 0)I(y_i^* > 0) + I(d_i = 1)I(d_i^* > 0)I(y_i^* \leq 0) + \\
I(d_i = 1)I(d_i^* > 0)I(y_{i0}^* > 0) + I(d_i = 1)I(d_i^* > 0)I(y_{i0}^* \leq 0)
\end{bmatrix}
\end{equation}

where

\[ s_i = \begin{bmatrix} d_i^* \\
d_i y_{i1}^* + (1-d_i) y_{i1}^{\text{miss}} \\
(C_i^*, y_{i0}^*) \end{bmatrix}, \quad r_i = \begin{bmatrix} W_i & 0 & 0 \\
0 & X_i & 0 \\
0 & 0 & X_i \end{bmatrix}, \quad \beta = \begin{bmatrix} \gamma \\
\beta_1 \\
\beta_0 \end{bmatrix}, \]

\( \phi_3(x; \mu, \Omega) \) denotes a trivariate normal density with mean \( \mu = [W\gamma, X\beta_1, X\beta_0] \), and covariance matrix \( \Omega \).

I specify priors of the form

\begin{align}
\beta &\sim N(\mu_\beta, V_\beta) \\
\rho_{D1} &\sim N(\mu_{D1}, V_{D1}) \\
\rho_{D0} &\sim N(\mu_{D0}, V_{D0}) \\
\rho_{10} &\sim N(\mu_{10}, V_{10})
\end{align}

with diffused values

\[
\mu_\beta = 0_{k\nu+2k \times 1}, V_\beta = 10^3 I_{k\nu+2k}, \\
\mu_{D1} = \mu_{D1} = \mu_{10} = 0, \quad V_{D1} = V_{D0} = V_{10} = 10^3.
\]

I provide full details on the posterior simulator in Appendix B along with a data generated simulation to demonstrate the performance of the posterior simulator. The sampling is performed using a Markov Chain Monte Carlo algorithm and cycles through the steps until convergence. The post convergence draws are used to produce the posterior distribution.

4.3. Treatment Effects

In this section I follow Poirier and Tobias (2003) to derive expression for conventional treatment effects for the outcome responses model. In particular, I adapt a conventional

---

15 Conditional on the latent variables, the model is a seemingly unrelated regressions (SUR) model except for the restriction that (1, 1) element of the covariance matrix is fixed at one.
treatment parameter of the Average Treatment Effect (ATE) to the response model, and
describe how it can be calculated within this framework.

The ATE typically quantifies the expected outcome gain for a randomly selected
individual. Given the observed data $D_i$ and $y_i$, I define the individual level gain from
receiving treatment as

$$
\Delta_i = y_i^* - y_i^0.
$$

(19)

Given definition (19), I characterize the following sampling distribution:

$$
[ATE]: p(\Delta \mid X, \Gamma)
$$

(20)

I am interested in characterizing the posterior predictive distributions of the outcome
gains. That is, I integrate out the parameters $\Gamma$ from the densities (20) by averaging over
the appropriate posterior distribution of those parameters. For the ATE, I would like to
obtain

$$
p(\Delta \mid X, \text{data}) = \int_{\Gamma} p(\Delta \mid X, \Gamma) p(\Gamma \mid \text{data}) d\Gamma
$$

To characterize the densities in (20), I define the following parameter:

$$
\delta_1 = 2(1 - \sigma_{i0}).
$$

Given the joint normality assumption, the following expression for (20) can be derived as
follows

$$
[ATE]: p(\Delta \mid X, \Gamma) = p_n(\Delta; X(\beta_1 - \beta_0), \delta_1)
= \delta_1^{-\frac{1}{2}} \phi\left( \frac{\Delta - X(\beta_1 - \beta_0)}{\sqrt{\delta_1}} \right)
$$

where $\Delta$ has a normal distribution with mean $X(\beta_1 - \beta_0)$ and variance $\delta_1$. Figure 2
depicts an individual’s treatment effect measured in the probability to exit teaching from
a unionized school and from a non-unionized school. In the figure, both $p(y_i^*)$ and
$p(y_{i0}^*)$ are less than .5, which implies that this individual is more likely to remain in
teaching in both unionized and non-unionized schools. $p(y_i^*) < p(y_{i0}^*)$ implies that this
individual is more likely to remain teaching in unionized schools than in non-unionized
schools.
5. Results

I run the Gibbs Sampler for 60,000 iterations and discard the first 10,000 of these as the burn-ins and construct the posterior distribution from every 50\textsuperscript{th} draw. Convergence is tested using Geweke’s Separated Partial Means (GSPM) test. For all sets of coefficients, I provide the mean posterior estimate on each coefficient along with the posterior standard deviation for the estimate. I also provide the probability of being greater than 0 for all coefficients. This gives an indication of how far from 0 the coefficient actually is.

Table 9 displays the marginal effects for the treatment equation: those who enter a teachers’ union. I highlight the most interesting results. First, teachers who earn their undergraduate degree in a state that have laws requiring collective bargaining in public schools are 15\% more likely to enter a union then state without such laws. Secondly, those who participate in a teacher practicum are 20\% more likely to enter union. This may be due to school of education graduates receiving practicum prior to entering teaching in a unionized school. It may also be due to teachers gaining in-school experience which makes them tend to favor unions. Third, those in households with income higher than $50,000 are 8\% more likely to enter teacher unions. Lastly, individuals’ socioeconomic characteristics do not seem to influence an individual’s decision to enter a union.
Table 9: Marginal Effects From Treatment Equation (Joining a Union) and Correlation Estimates

<table>
<thead>
<tr>
<th>Variable</th>
<th>Post mean</th>
<th>Post std.</th>
<th>Pr(• &gt; 0</th>
<th>Data)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-1.018</td>
<td>0.222</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>-0.020</td>
<td>0.041</td>
<td>0.26</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>-0.002</td>
<td>0.003</td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td>Non-White</td>
<td>0.025</td>
<td>0.045</td>
<td>0.59</td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>-0.001</td>
<td>0.040</td>
<td>0.40</td>
<td></td>
</tr>
<tr>
<td>ACT Score</td>
<td>-0.001</td>
<td>0.008</td>
<td>0.37</td>
<td></td>
</tr>
<tr>
<td>BA in Education</td>
<td>0.027</td>
<td>0.043</td>
<td>0.61</td>
<td></td>
</tr>
<tr>
<td>BA in Math / Science</td>
<td>0.089</td>
<td>0.070</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>Advanced Degree</td>
<td>-0.066</td>
<td>0.054</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>Not In Labor Force Last Year</td>
<td>0.043</td>
<td>0.037</td>
<td>0.74</td>
<td></td>
</tr>
<tr>
<td>Dependents under 5</td>
<td>-0.053</td>
<td>0.035</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>Household Income$^1$</td>
<td>0.082</td>
<td>0.037</td>
<td>0.82</td>
<td></td>
</tr>
<tr>
<td>Outside Teaching Job</td>
<td>0.057</td>
<td>0.067</td>
<td>0.68</td>
<td></td>
</tr>
<tr>
<td>Outside Non-Teaching Job</td>
<td>0.019</td>
<td>0.049</td>
<td>0.55</td>
<td></td>
</tr>
<tr>
<td>Class Size</td>
<td>0.010</td>
<td>0.004</td>
<td>0.83</td>
<td></td>
</tr>
<tr>
<td>Teacher Practicum</td>
<td>0.196</td>
<td>0.053</td>
<td>0.83</td>
<td></td>
</tr>
<tr>
<td>Special Education Assignment</td>
<td>-0.128</td>
<td>0.061</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>Math &amp; Science Assignment</td>
<td>-0.100</td>
<td>0.063</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>School Related Earnings$^{11}$</td>
<td>0.022</td>
<td>0.003</td>
<td>0.83</td>
<td></td>
</tr>
<tr>
<td>Charter School</td>
<td>-0.581</td>
<td>0.061</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Religious</td>
<td>-0.643</td>
<td>0.063</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Elementary School</td>
<td>0.028</td>
<td>0.042</td>
<td>0.63</td>
<td></td>
</tr>
<tr>
<td>Middle School</td>
<td>0.185</td>
<td>0.058</td>
<td>0.83</td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>0.033</td>
<td>0.045</td>
<td>0.64</td>
<td></td>
</tr>
<tr>
<td>% with BA or higher in community</td>
<td>-0.006</td>
<td>0.002</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Median Household Income in community$^{11}$</td>
<td>0.003</td>
<td>0.001</td>
<td>0.81</td>
<td></td>
</tr>
<tr>
<td>State Requires Collective Bargaining</td>
<td>0.147</td>
<td>0.038</td>
<td>0.83</td>
<td></td>
</tr>
</tbody>
</table>

Correlations

<table>
<thead>
<tr>
<th>$\rho_{\text{D1}}$</th>
<th>0.037</th>
<th>0.043</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_{\text{D0}}$</td>
<td>0.035</td>
<td>0.049</td>
</tr>
<tr>
<td>$\rho_{\text{10}}$</td>
<td>0.031</td>
<td>0.045</td>
</tr>
</tbody>
</table>
Table 10 contains the marginal effects for the potential outcomes equations. There are no individual physical or family structure characteristics that pinpoint which teachers are more likely to leave. Age, gender, race, marital status, and scholastic aptitude do not explain which teachers are more likely to exit teaching. Non-unionized teachers with a degree in education are twice as likely to remain in teaching than unionized teachers with the same degree. Similar results were found for math, science, and advanced degrees. Since this is an analysis of first year teachers, many do not have advanced degrees or master’s degrees. Unionized teachers who were previously unemployed are 3.9% more likely to exit teaching compared to the non-unionized teachers who are 4.9% less likely to exit teaching. This may reflect the ease in which unionized teachers can re-enter a unionized teaching position compared to a non-unionized position. Unionized teachers who were attacked or threatened the previous year are 10% more likely to leave teaching compared to 3.6% of non-unionized teachers. This reflects the expulsion tolerance of
school administrators to maintain teachers’ safety. Unionized teachers who had a teacher practicum were 7.9% less likely to leave teaching compared to 14.2% of non-unionized teachers. Interestingly, an increase of $1000 lowered the likelihood of leaving by only .6% for unionized teachers and .7% for non-unionized teachers. Thus, compared to a mandatory teacher practicum, raising teacher salary will do little to lower teacher attrition.

*Figure 3* displays individuals’ average latent propensities to exit teaching in unionized and non-unionized schools. The star points represent teachers who are observed in the data to teach in unionized schools while the plus points represent teachers who are observed to teach in non-unionized schools. 86.2% of all teachers have negative propensities to exit teaching in unionized schools compared to 72.5% of all teachers for non-union schools. Individuals’ propensities to exit teaching are much larger for teaching in unionized schools than non-unionized schools, especially for the negative propensities. This is not indicative of individuals having larger propensities to remain in teaching for unionized schools than non-unionized schools since the distributions are not identical. Calculating the propensities in probability provide a units of measure which are comparable. However, statement on individual’s consistency of choice across union status can be made. From this estimation, 91% of individuals are consistent in their exit decision across union status, i.e., their propensities to exit teaching in unionized and non-unionized schools are of the same sign. The remaining 9% have positive propensity for one union status and negative for the other.
The majority of these individuals, 68%, have negative propensities to exit for teaching in unionized schools and positive status for teaching in non-unionized schools, i.e., they exit teaching in non-unionized schools while remaining teaching in unionized schools.

*Figure 4* displays the cumulative distribution of the average treatment effect calculated as the difference in propensities to exit teaching between unionized and non-unionized schools which is measure in probability to exit teaching. A positive (negative) difference in probabilities indicates that the probability to exit teaching in a unionized school is greater (less) than the probability to exit teaching in non-union school. *Figure 4* indicates that unions have a negative effect on teacher attrition for 89% of all teachers. Of those negatively affected, the average effect is a 14.7% decrease in the probability of exiting teaching in a unionized school. Of those positively affected, the average effect is a 4.7% increase in the probability of exiting teaching in a union.
Table 11 displays the average treatment effect measured in the difference in probabilities to exit teaching in unionized and non-unionized schools. This is calculated explicitly since the counterfactual propensity to exit teaching is available from the posterior. For the sample of teachers, the ATE indicates that, on average, a randomly selected teacher is 16.7% more likely to remain in teaching due to unionization. The posterior standard deviation for this estimate is 4.3%.

| Table 11: ATE from Unionization on Teacher Attrition |
|---------------------------------|--------------|-----------------|-----------------|
| Average Treatment Effect        | Post Mean    | Post std.       | Pr(• > 0 | Data)       |
| Average Treatment Effect        | 0.167        | 0.043           | 1.00            |

The ATE is very far from 0. To calculate how much of the propensity to remain in teaching to attribute to the dismissal barriers unions possess, I divide each teachers’ treatment effect in probabilities through by their probability of remaining in teaching in unionized schools. This provides how much of the probability to remain teaching in a unionized school is attributed to the union. Table 12 provides the posterior predictive mean and standard deviation for a randomly selected teacher.
Table 12: The Effect of Dismissal Barriers on Teacher Attrition

<table>
<thead>
<tr>
<th>Attributed to Dismissal Barriers</th>
<th>Post Mean</th>
<th>Post std.</th>
<th>Pr((* &gt; 0 \mid Data))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.206</td>
<td>0.064</td>
<td>1.00</td>
</tr>
</tbody>
</table>

For the randomly selected teacher, on average, 20.6% of the propensity to remain in teaching is attributed to unionized while 79.4% is attributed to the type of individual who prepares for and teaches in unionized schools. Using the propensities of all teachers, I find that 3% of the 8% difference in attrition rates can be attributed to the union while 5% of the difference in attrition rate can be attributed to the self-selection into teaching. Thus, unions cannot account the majority of the difference found in the data. For unions to explain this much attrition, they would have to substantially increase the propensity to remain in teaching.

6. Conclusion

The purpose of this paper is to investigate the degree to which dismissal barriers affect teacher attrition. Teacher attrition is also affected by the self-selection of individuals into teaching. By explicitly calculating each teacher’s propensity to remain teaching in both unionized and non-unionized schools, the effect of unions on attrition is observable. For the majority of teachers, the decision to remain or stay is independent of union status. That is, most will exit or remain teaching regardless of whether they teach in a unionized school or not. This provides evidence that the difference in teacher attrition observed in the data has more to do with individuals’ self-selection process into teaching than to the barriers to dismissal unions provide. Thus, how one self-selects into teaching determines how long one stays in the teaching profession.

The teacher attrition literature has focused primarily on determining the influences on mobility decisions of teachers in public unionized schools. This paper’s contribution is to extend these analyses to understand the role of unions on teachers’ mobility decisions. While unions do provide many barriers to dismissal which lower the attrition rate for teachers in unionized schools, the magnitude is not as great as indicated in the data. I find that teachers’ unions can account for 21% of the propensity to remain in
teaching while 79% of the propensity to remain in teaching is due to individuals’ characteristics. While this may seem large, the estimates indicate that not many teachers are significantly influenced in changing their exit decision across union status. That is, 91% of all teachers are consistent with their exit decision across union status, either remaining or exiting in both teaching in unionized and non-unionized schools. 6% of all teachers would stay in a unionized school while exiting non-unionized schools while 3% of all teachers would stay in non-unionized schools while exiting unionized schools. Furthermore, only 3% of the 8% difference in attrition is attributed to the union while the remaining 5% is due to self-selection of teachers in unionized schools.

This work is a part of a larger research agenda. Gilpin and Kaganovich (2008) analyze the dynamical trade-off of quantity and quality of teachers over the last forty years for the United States. Collective bargaining contracts bind school administrators to salary schedules instead of market wages. The implication is that administrators face a trade-off of quantity and quality of teachers, and the bias shifts in favor of quantity given the evolution the wage distribution in the production economy exhibits an increasing skill premium. This tendency to prefer quantity to quality of teachers should cause the attrition rates to fall over time for those schools which have collective bargaining contracts. Table 12 provides the attrition rates for both unionized and non-unionized teacher attrition rates over the last 17 year for the U.S. Unionized attrition has fallen by 2% over this time period while non-unionized teacher attrition has risen. The analysis of Gilpin and Kaganovich (2008) demonstrates that with rising wage dispersion in the production economy over time, the optimal hiring policies of school administrators will be to raise the quantity of teachers at the expense of their relative quality. This would raise the barriers to dismissal as school administrators prefer to have a large quantity of teachers over their quality. While this result is consistent with facts, one can also observe at the same time, in non-unionized schools, the teacher attrition rate has fallen. One explanation is that it is difficult to retain individuals with higher outside of teaching wage rates.

---

16 In the 2004 data, the teacher attrition rate in non-unionized schools falls due to the large quantity of newly created charters which characteristically have low attrition rates.
Researchers have found mixed results from policies directed at raising the quantity and quality of teachers. Angrist and Lavy (1999) find that reducing class size causes significant and substantial increase in test scores for some grades but not others. Krueger (1999) analyzes data from Tennessee Project STAR finds that students’ scores increase by four percentage points in the first year of students attending smaller classes, while in subsequent years the test scores grow by about one percentage point per year. Hanushek (1999) rebuts Krueger’s findings citing important design and implementation issues from the STAR project that suggest the returns to class size reduction are biased upwards. Clotfelter et al. (2007) conclude that teacher experience, test scores and regular licensure all have greater positive effects on student achievement, whether compared to the effects of changes in class size or to the socioeconomic characteristics of students. In comparison Goldhaber and Anthony (2007), using the same data, find mixed evidence that improved observable teacher credentials have positive impact on student achievement. Rivkin et al. (2005) suggest that a ten student reduction in class size produces smaller benefits than one standard deviation improvement in teacher quality. Aaronson et al. (2007) find that replacing a teacher with another that is rated two standard deviations superior in quality can add 0.35 to 0.45 grade equivalents (or 30 to 40 percent of an average school year’s worth) to a student's math score performance.

Many education policies that are designed to increase either quantity or quality of teachers do receive an increase in education expenditures to fulfill these mandates. This forces school administrators to trade-off quantity or quality of teachers to satisfy the
mandate. One example is California’s class size reduction program. This program came at a cost of hiring lower quality teachers to staff additional classrooms, which offset the benefits of smaller classes (see Jepson and Rivkin (2002)). By not increasing expenditures, the type of teacher hired is different from the composition of current teachers. These teachers self-select into teaching for different reasons than current teachers and will remain in teaching for different durations. Future work is needed to understand how this change in composition of teachers has affected the teacher attrition rate over time. By estimating how self-selection and the effect of unionization have changed over time, researchers can obtain a better understanding on how current policies affect the self-selection process and how the barriers to dismissal affect differing compositions of entering teachers. This will also serve as an empirical validation to the quantity-quality trade-off as theorized by Gilpin and Kaganovich (2008).
References


Electronic References

Bureau of Economic Analysis (1): Current Dollar and Real GDP Data,  

Documentation for the 2004-05 Teacher Follow-up Survey, NCES product code  
2007349 (only available online),  

The National Center for Education Statistics: Schools and Staffing Survey:  
Restricted-access data, 2003-04 and 1999-01.

The National Center for Education Statistics: Teacher Follow-up Survey: Restricted- 
access data, 2004-05 and 2000-01.

The National Council on Teacher Quality: Legality of Collective Bargaining,  

U.S. Bureau of the Census (3): Summary File 3,  

U.S. Bureau of the Census: NAICS 2002 Occupational Codes  

University of California, San Diego: ACT to SAT converter,  
Appendix A: Data Construction

Two rounds of SASS (1999-00 and 2003-04) and two rounds of TFS (2000-01 and 2004-05) are pooled together to form one dataset. Pooling seems reasonable as the two rounds are close together in years and collective bargaining legislation has not changed significantly over the last six years.

The TFS is a stratified sample and has weights to aggregate the sample to the U.S. teaching population. According to the Documentation for the 2004-05 Teacher Follow-up Survey, the TFS weights are equal to the inverse of the sampled teacher’s probability of selection after three adjustments. The first adjustment was applied to reflect the impact of the SASS teacher weighting procedure. Next, a nonresponse adjustment factor was calculated and applied using information known about the respondents from the sampling frame data. Finally, a ratio adjustment factor was calculated and applied to the sample to adjust the sample totals to frame totals in order to reduce sampling variability.

Stratification can cause potential bias depending on the outcome of analysis. Cameron and Trivedi (2005) note that weighted estimators that adjust for differences in sampling rates may be necessary if the goal of analysis is a prediction of the population behavior. Such weighting is unnecessary if the conditional model for $y$ given $x$ is correctly specified and stratification is not on the dependent variable. If samples are determined in part by the value of the dependent variables, then weighted estimation is necessary. My primary interest lies with first year teachers with the comparison of unionized versus non-unionized attrition rates. The SASS was constructed to assure a minimum teacher sampling of 2,300 new teachers per school type (public, private). This constraint poses no problem for new public school teachers due to the large number of sampled schools with new teachers. However, for private school teachers, new teachers were oversampled by a factor of 1.5.\footnote{See Documentation for the 2004-05 Teacher Follow-up Survey pg. 20.} Since all private schools do not have teachers’ unions, without weighting the sample, this will lead to incorrect estimates of population impacts. For each population descriptive, I use the variable $tfsfinwt$ to weight the observations. Although I estimate the parameters using unweighted estimation, weighting
is used in subsequent impact calculations to predict population impacts, rather than sample impacts (See Cameron and Trivedi (2005) pg. 820-821).

Each teacher in the SASS is given an identification number. For teachers surveyed in both the SASS and the TFS, this identification number links the two surveys’ data at the teacher level. TFS do not attempt to link teachers across rounds which only permits repeated cross-sectional analysis to be conducted. This type of survey design creates a stock sample versus a flow sample used in duration analysis (See Lancaster, 1990). The main feature of stock sampling is spell length bias as the probability of being in the sample increases with spell length. The probability increases due to left truncation, i.e., those who have completed spells before the sample are omitted which causes those remaining in the sample to be seen with higher frequency. Stock sampling also has right censoring at the time of the sample since there are still teachers who haven’t yet teaching. Jenkins (1995) demonstrates a method to reformat the data that addresses both the left truncation and right censoring. This entails changing the unit of analysis from the individual to the time at risk of an event: exiting teaching. The main analysis focuses on first-year teachers where stock sampling isn’t an issue since the stock is also the flow.

The TFS contains most of the teacher and school characteristics needed in the analysis but I use other data sources to complete the dataset. The first ancillary data is from the 2000 Census which contains community-level data for each 5-digit zip code across the U.S. I use the zip code level data as private schools are not identifiable at the district level in the SASS/TFS surveys. The census data contains measures of socioeconomic well-being in the communities where teachers teach. The second ancillary data is from five rounds of National Postsecondary Student Aid Study (NPSAS) data. The TFS data set does not contain a measure of teacher aptitude, so ACT scores are imputed for each teacher equal to their undergraduate university’s ACT average. Second, the imputed teachers’ ACT scores are constructed using ACT scores and SAT scores converted to ACT scores at the university level.\footnote{It was brought to my attention that the ACT and SAT scoring have been adjusted over the last 20 years. This is to my advantage as I don’t have to deflate scores while aggregating them over time.} 180,670 student records are combined to create 2,052 university level enhanced ACT scores. There are 63.68 student records,
on average, per university to construct the enhanced ACT scores. The aggregation of individual ACT scores to university-level ACT scores flattens the tails of the distribution which can be seen with a maximum enhanced ACT score of 35.

The last ancillary data is from the National Council on Teacher Quality’s TR3 database on collective bargaining state laws. This permits an investigation on how state laws affect individuals’ decisions to enter teachers’ unions. The state level collective bargaining laws are matched to teachers by the state in which a teacher earned his/her undergraduate degree. These laws are then used as an instrument to identify the probability of joining a union but should not affect the mobility decision of teachers,

I now turn to the restriction I impose on the dataset. The first restriction is on class size which is restricted to less than 50 students per teacher, omitting one teacher with an average class size of 246. I also restrict teachers’ salaries to be above $10,000. This omits 4% of the sample. Lastly, age is restricted to be less than 50. This is somewhat arbitrary but necessary to minimize the retirement influence on teachers.

---

19 See Table C2 in Appendix C for a full listing on state collective bargaining laws.
### Table A1: Variable Labels, Definitions, and Data Sources

<table>
<thead>
<tr>
<th>Variable Label</th>
<th>Variable Description</th>
<th>Metric</th>
<th>2003-04 SASS &amp; 2004-05 TFS Variables</th>
<th>1999-00 SASS Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent Variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Member of teachers’ union</td>
<td>Individual reports member of teachers’ union</td>
<td>Binary</td>
<td>UNION_S</td>
<td>UNION</td>
</tr>
<tr>
<td>Exit Decision</td>
<td>Individual reports exiting teaching</td>
<td>Binary</td>
<td>STTUS_TF</td>
<td>STATUS</td>
</tr>
<tr>
<td><strong>Independent Variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>Individual reports being female</td>
<td>Binary</td>
<td>GENDER_S</td>
<td>GENDER</td>
</tr>
<tr>
<td>Age</td>
<td>Age on TFS survey</td>
<td>Z**</td>
<td>AGE_TF</td>
<td>AGE_TT</td>
</tr>
<tr>
<td>Non-White</td>
<td>Individual reports being non-white</td>
<td>Binary</td>
<td>RACETH_T</td>
<td>RACE</td>
</tr>
<tr>
<td>Married</td>
<td>Individual reports being married</td>
<td>Binary</td>
<td>F0234</td>
<td>F0197</td>
</tr>
<tr>
<td>Dependent under 5</td>
<td>Number of dependents under 5 on TFS</td>
<td>Z**</td>
<td>F0233</td>
<td>F0197</td>
</tr>
<tr>
<td>Outside Teaching Job</td>
<td>Outside teaching job during teaching year</td>
<td>Binary</td>
<td>T0406</td>
<td>T0354</td>
</tr>
<tr>
<td>Outside Non-Teaching Job</td>
<td>Outside non-teaching job during teaching year</td>
<td>Binary</td>
<td>T0406</td>
<td>T0354</td>
</tr>
<tr>
<td>Household Income1</td>
<td>Household income greater than 50K</td>
<td>Binary</td>
<td>F0231</td>
<td>F0194</td>
</tr>
<tr>
<td>Class Size</td>
<td>Number of students per teacher at school</td>
<td>Z**</td>
<td>STU_TCH</td>
<td>STU_TCH</td>
</tr>
<tr>
<td>Attacked or Threatened</td>
<td>Individual reports being attacked/threatened during last year</td>
<td>Binary</td>
<td>T0388 &amp; T0385</td>
<td>ATTACK &amp; THREAT</td>
</tr>
<tr>
<td>Teacher Practicum</td>
<td>Individual had a teacher practicum</td>
<td>Binary</td>
<td>T0155</td>
<td>T0128</td>
</tr>
<tr>
<td>Special Education Assignment</td>
<td>Individual's main assignment is sp. ed.</td>
<td>Binary</td>
<td>ASSIGN03</td>
<td>ASSIGN</td>
</tr>
<tr>
<td>Math &amp; Science Assignment</td>
<td>Individual's main assignment is math or sci.</td>
<td>Binary</td>
<td>ASSIGN03</td>
<td>ASSIGN</td>
</tr>
<tr>
<td>Not In Labor Force Last Year</td>
<td>Individual not in labor force prior to teaching</td>
<td>Binary</td>
<td>T0030</td>
<td>T0059</td>
</tr>
<tr>
<td>BA in Education</td>
<td>Individual reports a BA major in education</td>
<td>Binary</td>
<td>T0119, T0121, and T0132</td>
<td>T0072, T0074, and T0088</td>
</tr>
<tr>
<td>BA in Math / Science</td>
<td>Individual reports a BA major in math/science</td>
<td>Binary</td>
<td>T0119, T0121, and T0132</td>
<td>T0072, T0074, and T0088</td>
</tr>
<tr>
<td>Advanced Degree</td>
<td>Individual has an advanced degree</td>
<td>Binary</td>
<td>HIDEGR_S</td>
<td>HIDEGREE</td>
</tr>
<tr>
<td>School Related Earnings11</td>
<td>Individuals reported school related earnings</td>
<td>Z**</td>
<td>EARNSC</td>
<td>EARNSC</td>
</tr>
<tr>
<td>Charter School</td>
<td>Individual teaches at a charter school</td>
<td>Binary</td>
<td>CHARFLAG</td>
<td>SECTOR</td>
</tr>
<tr>
<td>Religious</td>
<td>Individual teaches at a religious school</td>
<td>Binary</td>
<td>RELIG</td>
<td>RELIG</td>
</tr>
<tr>
<td>Elementary School</td>
<td>Individual teaches at an elementary school</td>
<td>Binary</td>
<td>SCHLEVE2</td>
<td>SCHLEVE2</td>
</tr>
<tr>
<td>Middle School</td>
<td>Individual teaches at a middle school</td>
<td>Binary</td>
<td>SCHLEVE2</td>
<td>SCHLEVE2</td>
</tr>
<tr>
<td>Urban</td>
<td>Individual teaches in an urban community</td>
<td>Binary</td>
<td>URBAN03</td>
<td>URBANIC</td>
</tr>
<tr>
<td><strong>Ancillary Data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACT Score</td>
<td>Imputed ACT score</td>
<td>Z**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% with BA or higher</td>
<td>Percent of community with BA or adv. degree</td>
<td>Percent</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median Household Income1</td>
<td>Community's median household income</td>
<td>Z**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>State Requires Col. Bargaining</td>
<td>State requires collective bargaining</td>
<td>Binary</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Footnotes**

1: Recoded L to 1, S to 2, and M to 3. 7: Codes 101-136, 160-162, and 181-182.
2: Recoded 3 to 5. 8: Codes 01 - 44.
3: Recoded 1,2 to 1, 3,4 to 2, and 5,6 to 3. 9: Codes 190 - 218.
4: Recoded 1,2,3 to 1, 4,5 to 2, and 6 to 3. 10: Codes 57 - 63, and 68 - 69.
5: Recoded 2 to 1. 11: In 1000s of 2004 dollars.
6: Recoded 2 to 8, 3 to 4, 4 to 10, 5 to 2, 7 to 5, and 8 to 11. 12: Recoded D to 1.
<table>
<thead>
<tr>
<th>Variable Label</th>
<th>Data Source</th>
<th>Year</th>
<th>Variables</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACT Score¹</td>
<td>NPSAS</td>
<td>1990</td>
<td>ACT sattotal²</td>
<td>6316</td>
</tr>
<tr>
<td></td>
<td>NPSAS</td>
<td>1992</td>
<td>ACT sattotal²</td>
<td>23592</td>
</tr>
<tr>
<td></td>
<td>NPSAS</td>
<td>1996</td>
<td>ACT sattotal²</td>
<td>41256</td>
</tr>
<tr>
<td></td>
<td>NPSAS</td>
<td>2000</td>
<td>TEACTCRE TESATDER²</td>
<td>35481</td>
</tr>
<tr>
<td></td>
<td>NPSAS</td>
<td>2004</td>
<td>ASTACTS SAT²</td>
<td>24025</td>
</tr>
<tr>
<td>% with BA or higher³</td>
<td>U.S. Census Summary File 3 (SF 3)</td>
<td>2000</td>
<td>P037015 P037032</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2000</td>
<td>P037016 P037033</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2000</td>
<td>P037017 P037034</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2000</td>
<td>P037018 P037035</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2000</td>
<td>P037001</td>
<td></td>
</tr>
<tr>
<td>Median Household Income³</td>
<td>U.S. Census Summary File 3 (SF 3)</td>
<td>2000</td>
<td>P053001</td>
<td></td>
</tr>
</tbody>
</table>

Footnotes
1: Linked to TFS by IPEDS number.
2: SAT are converted to ACT using UC-SD converter (see Electronic references)
3: Linked to TFS by 5-digit zip code.
## Table A3: Summary Statistics by Teachers' Association Status

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Union (Treated state)</th>
<th>Non-Union (Untreated state)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std</td>
</tr>
<tr>
<td><strong>Dependent Variables</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Member of Teachers’ Association</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Exit Decision</td>
<td>0.08</td>
<td>5.59</td>
</tr>
<tr>
<td><strong>Independent Variables</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>0.74</td>
<td>8.9</td>
</tr>
<tr>
<td>Age</td>
<td>29.9</td>
<td>153.4</td>
</tr>
<tr>
<td>Non-White</td>
<td>0.18</td>
<td>7.8</td>
</tr>
<tr>
<td>Married</td>
<td>0.46</td>
<td>10.2</td>
</tr>
<tr>
<td>Dependents under 5</td>
<td>0.28</td>
<td>11.4</td>
</tr>
<tr>
<td>Outside Teaching Job</td>
<td>0.05</td>
<td>4.6</td>
</tr>
<tr>
<td>Outside Non-Teaching Job</td>
<td>0.15</td>
<td>7.3</td>
</tr>
<tr>
<td>Household Income(^1)</td>
<td>0.55</td>
<td>10.1</td>
</tr>
<tr>
<td>Class Size</td>
<td>15.6</td>
<td>83.0</td>
</tr>
<tr>
<td>Attacked or Threatened</td>
<td>0.10</td>
<td>6.1</td>
</tr>
<tr>
<td>Teacher Practicum</td>
<td>0.92</td>
<td>5.7</td>
</tr>
<tr>
<td>Special Education Assignment</td>
<td>0.10</td>
<td>6.1</td>
</tr>
<tr>
<td>Math &amp; Science Assignment</td>
<td>0.16</td>
<td>7.4</td>
</tr>
<tr>
<td>ACT Score</td>
<td>21.5</td>
<td>46.6</td>
</tr>
<tr>
<td>Not In Labor Force Last Year</td>
<td>0.48</td>
<td>10.2</td>
</tr>
<tr>
<td>BA in Education</td>
<td>0.56</td>
<td>10.1</td>
</tr>
<tr>
<td>BA in Math / Science</td>
<td>0.11</td>
<td>6.5</td>
</tr>
<tr>
<td>Advanced Degree</td>
<td>0.16</td>
<td>7.4</td>
</tr>
<tr>
<td>School Related Earnings(^2)</td>
<td>32.8</td>
<td>126.5</td>
</tr>
<tr>
<td>Charter School</td>
<td>0.01</td>
<td>1.5</td>
</tr>
<tr>
<td>Religious</td>
<td>0.01</td>
<td>1.6</td>
</tr>
<tr>
<td>Elementary School</td>
<td>0.51</td>
<td>10.2</td>
</tr>
<tr>
<td>Middle School</td>
<td>0.19</td>
<td>8.0</td>
</tr>
<tr>
<td>Urban</td>
<td>0.86</td>
<td>7.0</td>
</tr>
<tr>
<td>% with BA or higher</td>
<td>22.1</td>
<td>276.2</td>
</tr>
<tr>
<td>Median Household Income(^2)</td>
<td>55.9</td>
<td>428.5</td>
</tr>
</tbody>
</table>

**Instrument on decision to join a teachers’ association**

<table>
<thead>
<tr>
<th></th>
<th>Union (Treated state)</th>
<th>Non-Union (Untreated state)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>obs = 469</td>
<td>obs = 615</td>
</tr>
<tr>
<td>State Requires Collective Bargaining</td>
<td>0.77</td>
<td>0.73</td>
</tr>
<tr>
<td>Where Teacher Earned Bachelor's Degree</td>
<td>8.6</td>
<td>8</td>
</tr>
</tbody>
</table>


**Footnotes**

1: Above $50,000.
2: In 1000s of 2004 dollars.
Appendix B: The Posterior Simulator and Data Generated Experiment

B.1. The Posterior Simulator

Blow is the posterior simulator for fitting the parameterized treatment-response model. I adopt the standard notation $\Gamma_{-x}$ to denote all parameters other than $x$. I first group the joint posterior into $[D^*, y_1^*, y_0^*, \beta, \Sigma]$. The latent data will be sampled in blocking steps, while the regression parameters and covariance matrix will be drawn from their complete posterior conditional. The Gibbs sampler cycles through steps 1 through 6 as described below until convergence is obtained. Post convergence draws are used to produce the posterior distribution.

MCMC Algorithm for sampling

Step 1: Sampling $\beta$

Conditioned on the values of the observed and missing latent data, it is straightforward to sample from the complete conditional

$$\beta | \Gamma_{-\beta}, y, D : N(D_{\beta}d_{\beta}, D_{\beta}), \tag{21}$$

where

$$D_{\beta} = [R'(\Sigma^{-1} \otimes I_n)R + V\beta^{-1}]^{-1} \text{ and } d_{\beta} = [R'(\Sigma^{-1} \otimes I_n)S + V\beta^{-1} \mu_{\beta}].$$

and

$$S_{3n+1} \equiv \begin{bmatrix} D^* \\ DY_1^* + (1-D)Y_1^{*miss} \\ DY_0^{*miss} + (1-D)Y_0^* \end{bmatrix}, \text{ and } R_{3n+(k_{p}+2k)} \equiv \begin{bmatrix} W & 0 & 0 \\ 0 & X & 0 \\ 0 & 0 & X \end{bmatrix}.$$

Step 2: Sample $\Sigma$

In sampling $\Sigma$, I exploit the fact that the error vectors are already known. That is, one can sample from the complete conditionals given:

$$\sigma_{p1} | \Gamma_{-\sigma_{p1}}, y^*, y, D : N(D_{p1}d_{p1}, D_{p1})$$

$$\sigma_{p0} | \Gamma_{-\sigma_{p0}}, y^*, y, D : N(D_{p0}d_{p0}, D_{p0})$$

$$\sigma_{10} | \Gamma_{-\sigma_{10}}, y^*, y, D : N(D_{10}d_{10}, D_{10}) \tag{22}$$
where

\[
D_{D1} = [\dot{e}_D e_D + V_{D1}\mathbf{1}]^{-1} \quad \text{and} \quad d_{D1} = [\dot{e}_D e_D + V_{D1}^{-1}\mu_{D1}]
\]

\[
D_{D0} = [\dot{e}_D e_D + V_{D0}\mathbf{1}]^{-1} \quad \text{and} \quad d_{D0} = [\dot{e}_D e_D + V_{D0}^{-1}\mu_{D0}]
\]

\[
D_{i0} = [\dot{e}_i e_i + V_{i0}]^{-1} \quad \text{and} \quad d_{i0} = [\dot{e}_i e_i + V_{i0}^{-1}\mu_{D1}].
\]

I reject \(\sigma_{D1}, \sigma_{D0},\) and \(\sigma_{i0}\) draws that are not between \([0,1]\) so as to have proper interpretation as covariance/correlations. From these draws, I construct \(\Sigma\). Since, as noted early, \(\Sigma\) must be positive definite, I accept/reject \(\Sigma\) based on all eigenvalues being negative.\(^{20}\)

The remaining steps in the posterior simulator involve sampling of the latent data \(S\). Since the Jacobian of the transformation from the error vector to the outcome quantities is unity, the trivariate normal distribution for \([D^*_i, y^*_i, y^*_{i0}]\) is easily derived. Assuming observations are independent of one another, I can draw from univariate normals. I proceed by sampling from the following densities:

\[
y^*_i | y^*_0, D^*, \Gamma, y, D
\]

\[
y^*_i | y^*_0, D^*, \Gamma, y, D
\]

and

\[
D^* | y^*_i, y^*_0, \Gamma, y, D
\]

**Step 3: Sample observed and missing latent outcome data** \(y^*_{1n}\) and \(y^*_{10n}\) from the conditional

\[
y^*_{i1} | y^*_0, D^*, \Gamma, y, D \sim \begin{cases} \text{TN}_{(0,\infty)}(\mu_{i1}, \sigma_{i1}) & \text{if } d_{i1} = 1 \text{ & } y_{i1} = 1 \\ \text{TN}_{(-\infty,0]}(\mu_{i1}, \sigma_{i1}) & \text{if } d_{i1} = 1 \text{ & } y_{i1} = 0, \quad i = 1, 2, K, n \\ N(\mu_{i1}, \sigma_{i1}) & \text{if } d_{i1} = 0 \end{cases}
\]  

\(^{20}\) See Li (1998) for further discussion.

\(^{21}\) A Metropolis-Hastings algorithm inside of the Gibbs Sampler can be substituted in this step. During simulations, both performed equally well in terms of estimating the true parameters. In terms of speed, obviously the Metropolis-Hastings algorithm is much slower.
where
\[
\mu_1 \equiv X_1 \beta_1 + \left[\sigma_{D_1} \sigma_{10}\right]^{-1} \begin{bmatrix} 1 & \sigma_{D_0}^* \\ \sigma_{D_0} & 1 \end{bmatrix}^{-1} \begin{bmatrix} d_i^* - W_i \gamma^* \\ y_{10}^* - X_i \beta_0 \end{bmatrix}
\]

and
\[
\sigma_1 \equiv 1 - \left[\sigma_{D_1} \sigma_{10}\right]^{-1} \begin{bmatrix} 1 & \sigma_{D_0} \\ \sigma_{D_0} & 1 \end{bmatrix}^{-1} \begin{bmatrix} \sigma_{D_1} \\ \sigma_{10} \end{bmatrix}.
\]

\(\sigma_{D_1}, \sigma_{D_0}\), and \(\sigma_{10}\) refer to the covariance/correlation parameters between the treated and untreated outcome errors and outcome and selection equation errors. \(TN_{(a,b)}(\mu, \sigma)\) denotes a univariate normal density with mean \(\mu\) and variance \(\sigma\), truncated to the interval \((a,b)\).

With the missing or potential outcomes, I simply draw from a normal density since there are no restrictions. The observed outcomes are related to latent variables using standard textbook techniques of drawing from a truncated normal density.

**Step 4: Sample observed and missing latent outcome data** \(y_{0_{0\theta}}^*\) and \(y_{0_{0\theta}}^{\text{miss}}\) from the conditional

Similarly, I can sample from:
\[
y_{i0}^* \mid y_1^*, D^*, \Gamma, y, D \sim \begin{cases} TN_{(\mu_{i0}, \sigma_{i0})} & \text{if } d_i = 0 \text{ & } y_{i0} = 1 \\ TN_{(-\infty, 0]}(\mu_{i0}, \sigma_{i0}) & \text{if } d_i = 0 \text{ & } y_{i0} = 0, \quad i = 1, 2, K, n \end{cases} \tag{27}
\]

where
\[
\mu_{i0} \equiv X_i \beta_0 + \left[\sigma_{D_0} \sigma_{10}\right]^{-1} \begin{bmatrix} 1 & \sigma_{D_1} \\ \sigma_{D_1} & 1 \end{bmatrix}^{-1} \begin{bmatrix} d_i^* - W_i \gamma^* \\ y_{10}^* - X_i \beta_0 \end{bmatrix}
\]

and
\[
\sigma_{i0} \equiv 1 - \left[\sigma_{D_0} \sigma_{10}\right]^{-1} \begin{bmatrix} 1 & \sigma_{D_1} \\ \sigma_{D_1} & 1 \end{bmatrix}^{-1} \begin{bmatrix} \sigma_{D_0} \\ \sigma_{10} \end{bmatrix}.
\]
Step 5: Sample missing latent data $D^*$

The conditional distribution of the latent data $D^*_i$ is complicated by the fact that I know the sign. With this minor complication, I draw similarly to the latent outcomes from a conditional normal, which is truncated by the observed value of $D_i$:

$$
d^*_i \mid \Gamma - d^*_i, y, D \sim \begin{cases} 
TN_{(0, \infty)}(\mu_{id}, \sigma_{id}) & \text{if } d_i = 1, \\
TN_{(-\infty, 0)}(\mu_{id}, \sigma_{id}) & \text{if } d_i = 0
\end{cases} \quad i = 1, 2, \ldots, n \quad (28)
$$

where

$$
\mu_{id} = W_i \gamma + \begin{bmatrix} \sigma_{D1} \\ \sigma_{D0} \end{bmatrix} \begin{bmatrix} 1 & \sigma_{10} \\ \sigma_{10} & 1 \end{bmatrix}^{-1} \begin{bmatrix} y^*_i - X_i \beta_1 \\ y^*_i - X_i \beta_0 \end{bmatrix}
$$

and

$$
\sigma_{id} = 1 - \begin{bmatrix} \sigma_{D1} \\ \sigma_{D0} \end{bmatrix} \begin{bmatrix} 1 & \sigma_{10} \\ \sigma_{10} & 1 \end{bmatrix}^{-1} \begin{bmatrix} \sigma_{D1} \\ \sigma_{D0} \end{bmatrix}
$$

Step 6: Repeat 1-5 using the most recent values of the conditioning variables.
B.2. Data Generated Experiment

I conduct a generated data simulation to demonstrate the performance of the posterior simulator. A sample of $N = 5000$ observations is generated from the following potential outcomes model:

$$d_i^* = \gamma_0 + \gamma_1 w_{i1} + \gamma_2 w_{i2} + \epsilon_{iD},$$
$$y_{i1}^* = \beta_{10} + \beta_{11} x_{i1} + \epsilon_{i1},$$
$$y_{i0}^* = \beta_{00} + \beta_{01} x_{i1} + \epsilon_{i0},$$

where $w_{ik}$ is drawn independently from a $N(0,1)$ distribution $k = 1, 2$ as well as $x_{i1}$.

The error terms $[\epsilon_{iD} \epsilon_{i1} \epsilon_{i0}]'$ are drawn jointly from the trivariate Normal distribution

$$
\begin{bmatrix}
\epsilon_{iD} \\
\epsilon_{i1} \\
\epsilon_{i0}
\end{bmatrix}
\sim N
\begin{bmatrix}
0 \\
0 \\
0
\end{bmatrix}, \begin{bmatrix}
1 & .2 & .2 \\
.2 & 1 & .2 \\
.2 & .2 & 1
\end{bmatrix}.
$$

Finally, the regression parameters are set as follows

$$\gamma = [2.3, 5, 1], \quad \beta_1 = [1.2, 2], \quad \beta_0 = [.5, 1.5]$$

I follow the steps for 2000 draws, dropping the first 500 as burn-ins and keeping every 5th draw. Table 6 provides a comparison of the true values and the values obtained from the posterior distribution. Overall, the simulator works well in estimating the parameter values and the latent variables. To correct for autocorrelation, I take every 5th draw which seems to alleviate the issue. This is an inherent issue with limited dependent models in general. In terms of performance, all parameter estimates are within two standard deviations and all but two are within one standard deviation of the true values. As this experiment was conducted several times, performance remains consistent.
<table>
<thead>
<tr>
<th>Regression Parameters</th>
<th>True Value</th>
<th>Post Mean</th>
<th>Post Std.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma_0$</td>
<td>0.2</td>
<td>0.180</td>
<td>0.03</td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>3.5</td>
<td>3.546</td>
<td>0.10</td>
</tr>
<tr>
<td>$\gamma_2$</td>
<td>1</td>
<td>0.982</td>
<td>0.05</td>
</tr>
<tr>
<td>$\beta_{10}$</td>
<td>1.2</td>
<td>1.288</td>
<td>0.06</td>
</tr>
<tr>
<td>$\beta_{11}$</td>
<td>2</td>
<td>1.866</td>
<td>0.14</td>
</tr>
<tr>
<td>$\beta_{00}$</td>
<td>0.5</td>
<td>0.399</td>
<td>0.04</td>
</tr>
<tr>
<td>$\beta_{01}$</td>
<td>1.5</td>
<td>1.425</td>
<td>0.06</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Latent Effects</th>
<th>True Value</th>
<th>Post Mean</th>
<th>Post Std.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D^*$</td>
<td>0.2438</td>
<td>0.2166</td>
<td>0.0301</td>
</tr>
<tr>
<td>$Y^*_1$</td>
<td>1.2245</td>
<td>1.3111</td>
<td>0.0609</td>
</tr>
<tr>
<td>$Y^*_0$</td>
<td>0.5359</td>
<td>0.4154</td>
<td>0.0401</td>
</tr>
</tbody>
</table>

Footnotes
1: Results are based on 2,000 MCMC draws.
Appendix C. Associated Tables

<table>
<thead>
<tr>
<th>Code</th>
<th>Code Description</th>
<th>Occupation Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Management occupations</td>
<td>0010-0430</td>
</tr>
<tr>
<td>2</td>
<td>Business and financial operations occupations</td>
<td>0500-0950</td>
</tr>
<tr>
<td>3</td>
<td>Computer and mathematical science occupations</td>
<td>1000-1240</td>
</tr>
<tr>
<td>4</td>
<td>Architecture and engineering occupations</td>
<td>1300-1560</td>
</tr>
<tr>
<td>5</td>
<td>Life, physical, and social science occupations</td>
<td>1600-1960</td>
</tr>
<tr>
<td>6</td>
<td>Community and social service occupations</td>
<td>2000-2060</td>
</tr>
<tr>
<td>7</td>
<td>Legal occupations</td>
<td>2100-2150</td>
</tr>
<tr>
<td>8</td>
<td>Education, training, and library occupations</td>
<td>2200-2550</td>
</tr>
<tr>
<td>9</td>
<td>Arts, design, entertainment, sports, and media occupations</td>
<td>2600-2960</td>
</tr>
<tr>
<td>10</td>
<td>Healthcare practitioner and technical occupations</td>
<td>3000-3540</td>
</tr>
<tr>
<td>11</td>
<td>Healthcare support occupations</td>
<td>3600-3650</td>
</tr>
<tr>
<td>12</td>
<td>Protective service occupations</td>
<td>3700-3950</td>
</tr>
<tr>
<td>13</td>
<td>Food preparation and serving related occupations</td>
<td>4000-4160</td>
</tr>
<tr>
<td>14</td>
<td>Building and grounds cleaning and maintenance occupations</td>
<td>4200-4250</td>
</tr>
<tr>
<td>15</td>
<td>Personal care and service occupations</td>
<td>4300-4650</td>
</tr>
<tr>
<td>16</td>
<td>Sales and related occupations</td>
<td>4700-4960</td>
</tr>
<tr>
<td>17</td>
<td>Office and administrative support occupations</td>
<td>5000-5930</td>
</tr>
<tr>
<td>18</td>
<td>Farming, fishing, and forestry occupations</td>
<td>6000-6130</td>
</tr>
<tr>
<td>19</td>
<td>Construction and extraction occupations</td>
<td>6200-6940</td>
</tr>
<tr>
<td>20</td>
<td>Installation, maintenance, and repair occupations</td>
<td>7000-7620</td>
</tr>
<tr>
<td>21</td>
<td>Production occupations</td>
<td>7700-8960</td>
</tr>
<tr>
<td>22</td>
<td>Transportation and material moving occupations</td>
<td>9000-9750</td>
</tr>
<tr>
<td>23</td>
<td>Armed Forces</td>
<td>9840</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>State</th>
<th>Minimum number of years before tenure is granted</th>
<th>Legality of Collective Bargaining</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alabama</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Alaska</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Arizona</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Arkansas</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>California</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Colorado</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Connecticut</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Delaware</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Florida</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Georgia</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Hawaii</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Idaho</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Illinois</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Indiana</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Iowa</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Kansas</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Kentucky</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Louisiana</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Maine</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Maryland</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Massachusetts</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Michigan</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Minnesota</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Mississippi</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Missouri</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Montana</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Nebraska</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Nevada</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>New Hampshire</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>New Jersey</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>New Mexico</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>New York</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>North Carolina</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>North Dakota</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Ohio</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Oklahoma</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Oregon</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Pennsylvania</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Rhode Island</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>South Carolina</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>South Dakota</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Tennessee</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Texas</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Utah</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Vermont</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Virginia</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Washington</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>West Virginia</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Wisconsin</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Wyoming</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

Source: TR3 Database From The National Council on Teacher Quality

Footnote: 1 State requires collective bargaining. 2 Collective bargaining is permissible. 3 Collective bargaining is explicitly illegal.
Chapter 3
Unequal Longevity and Social Security:
The Implications for Human Capital, Growth and Distribution

Abstract

This paper analyzes the implications of pay-as-you-go (PAYG) social security system for human capital investment, economic growth and income distribution when longevity is dependent on human capital attainment. The results indicate that when a social security system is implemented, the effects on growth and welfare inequality are temporally non-monotonic. While PAYG pensions encourage human capital investment at a cost of lower growth in the short-run, it results in increased life expectancy of the subsequent generations, providing additional incentive to save for retirement. Thus, in the long-run, implementing PAYG social security causes the stock of human capital to grow; savings will increase due to the dynamic externality, thereby so will the aggregate output.

*JEL* classification: H55, D91, I28, O4
1. Introduction

It is well-known in the literature that pay-as-you-go (PAYG) social security systems may have a negative effect on physical capital accumulation due to a reduction of savings for retirement. The standard thought is that this is a trade-off to lowering welfare inequality. Recent research suggests that PAYG social security provides incentives for working-age individuals to invest in education and other human capital of the younger generation in order to increase pension benefits in retirement (See Bellettini and Certoni (1999), Kaganovich and Zilcha (1999), Glomm and Kaganovich (2008), and Kaganovich and Meier (2008)). Under certain conditions these investments can lead to higher growth while reducing inequality. This paper continues this research by exploring the dynamic externality generated by social security on subsequent generations, that of inadvertently increasing the younger generation’s longevity due to increasing their human capital.

Most of the theoretical research on the impact of social security systems on growth and welfare inequality assumes that all individuals have identical longevities (See Glomm and Kaganovich (2008), Kaganovich and Meier (2008), and Zhang, Zhang and Leung (2006)). Glomm and Kaganovich (2008) study the growth-inequality relationship assuming that individuals have identical longevities in an overlapping generations (OLG) model with PAYG Beveridgian social security. They find that if the social security system is sufficiently small, then growth can be enhanced while reducing welfare inequality. Zhang, Zhang and Leung (2006) explore the effect of annuity-based PAYG Beveridgian social security on output and welfare with homogenous agents. Zhang et al. (2006) find that aggregate welfare increases at a cost of lower output.22 Kaganovich and Meier (2008) study the growth implications of fully funded versus unfunded Beveridgian and Bismarckian social security schemes, assuming that individuals have identical mortality. They conclude that Beveridgian PAYG pensions achieve the highest growth rate relative to the other pension designs. In their paper, a small open economy is

---

22 This precludes an analysis on the progressivity of social security. See Echevarría and Iza (2006) as well.
considered which implies that there is no adverse effect on the physical capital stock from the PAYG design. Neither of these papers addresses welfare inequality.23

The purpose of this paper is to study the dynamical effects of PAYG social security schemes on short-run and long-run growth and welfare inequality. Up to this point, the results on growth and welfare tend to be temporally monotonic; welfare inequality and growth increase or decrease for all subsequent generations after a social security scheme is implemented. The argument in this paper is that implementing PAYG Beveridgian social security schemes raises the human capital of the younger generations which, in turn, increases longevity of subsequent generations. The increase in longevity raises individuals’ concern for retirement. Confronted with a longer retirement, individuals devote a higher fraction of their lifetime resources to savings at the cost of consumption and investments in children’s human capital. Thus, the dynamic externality of raising human capital of younger generations affects future generations’ human capital. This analysis investigates the effect of this dynamic externality on human capital accumulation and growth.

This work is closely related to another line of research which studies the effect of unequal longevity on the progressivity of social security. Hachon (2008) studies how varying the dependency of longevity on innate ability, which directly predicts earnings capabilities, affects the progressivity of various social security systems. He finds that when longevity is dependent on innate ability, implementing a purely Beveridgian social security system makes the pension benefit formula a function of longevity, reducing inequality only if the distribution of longevity is more equal than the distribution of human capital. Empirical research estimating the degree of progressivity for the U.S. social security system indicates that accounting for shorter longevities of poorer individuals, the U.S. social security system is not as progressive as previously perceived and may be slightly regressive (see Brown (2003), Coronado, Fullerton, and Glass (2000), Garrett (2007), Kotlikoff et al. (1998)), and Panis and Lillard (1996).

23 In Kaganovich and Meier (2008), it can be inferred that the Beveridgian formula reduces inequality while the Bismarckian formula does not.
I first explore the effect of implementing Beveridgian PAYG social security assuming that longevity is independent of human capital attainment to provide a benchmark. These results are standard in the literature. I then assume longevity is dependent on human capital attainment and study the effect of implementing social security. Using transition paths provides a basis of comparison for the effects in the short and long run. Assuming longevity is independent versus dependent of human capital attainment on intra-generational transfers and on short/long run growth and welfare inequality are quite distinct. Under either assumption, implementing PAYG social security acts as the vehicle to lowering welfare inequality. However, the effect on growth is conditional upon two factors: 1) dependency of longevity on human capital attainment and 2) the size of the social security system.

When longevity is independent of human capital attainment, the standard results found in the literature are confirmed: under certain parametric conditions, if the size of social security is sufficiently small, then implementing social security enhances growth. It is now well known in the literature that the PAYG feature provides working-age individuals incentives to invest in the human capital of the young generation in order to increase their social security benefits in retirement. Growth is enhanced as long as the increase in human capital is sufficiently large to offset the decrease in savings. As the social security system becomes large, savings decrease causing the task of compensating this loss in production to be increasingly more difficult.

When longevity is dependent on human capital attainment, many of the benchmark results may not hold. I find that when the social security system is sufficiently large, growth exhibits a temporally non-monotonic relationship: growth declines in the short-run and rises in the long-run. There are two effects from implementing social security which drive this result, one which is standard and the other which is a new insight provided by this paper. The first one is that the PAYG formula causes individuals to increase investments in the younger generation’s human capital. The Beveridgian formula causes human capital of the poor to increase more than the human capital of the wealthy. Thus, aggregate human capital of the younger generation rises while reducing welfare inequality. When social security is large, the increase in human capital is not
sufficient to compensate for the decrease in savings from the PAYG design, lowering output in the short-run.

The second effect is the dynamic externality of lengthening future generation’s longevity due to their increased human capital. Increasing longevity raises concern for consumption during retirement. Confronted with a longer retirement, individuals devote a higher fraction of their lifetime resources to savings. Thus, in the long-run, implementing PAYG social security causes the stock of human capital to grow; in the long-run savings will increase due to the dynamic externality, thereby so will the aggregate output.

The paper is structured as follows. Section 2 sets up the model, Section 3 defines the Dynamic Competitive Equilibrium, Section 4 presents the solution of the benchmark model, and Sections 5 provides the results for implementing social security schemes. Section 6 concludes. The appendices contain numerical results.

2. The Model
Consider an OLG economy populated by individuals who live for three periods, youth, working-age, and retirement. Each individual is indexed by family name \( \omega \) and by generation which is identified with the period the individual is working-age. Individuals born in period \( t - 1 \) make up generation \( t \). Let \( \Omega_t \) denote the set of individuals of generation \( t \) and \( \mu_t(\omega) \) be the Lebesgue measure which is assumed to be well defined. The population is constant and normalized to 1. Youth and working-age periods both have length 1 while the last period of life, retirement, has length \( T \leq 1 \).

Individuals have two economically active periods, working-age and retirement. During the first period of life, children remain economically inactive, making no decisions of their own, while acquiring human capital. Individuals become active during their working-age, deriving income from work, making consumption decisions, and investing in both savings and human capital of their children. In retirement, individuals finance consumption from the returns on savings and an annuity-based public pension, if available.
2.1. Human Capital Formation

The human capital of a young individual $\omega$ in generation $t+1$ is produced using a private input contributed by his parent, $e_t(\omega)$, in the production function

$$h_{t+1}(\omega) = Be_t(\omega)$$

(1)

where $B > 0$.

I assume that human capital is also the effective labor supply which implies that human capital attainment and earning capabilities are equivalent. The aggregate (and average) level of human capital of generation $t$ is

$$H_t = \int_{\Omega} h_t(\omega)d\mu_t(\omega)$$

(2)

The initial human capital endowments, $h_0(\omega)$ for all $\omega \in \Omega_0$, are assumed to be given.

2.2. Production

The production sector of the economy consists of private, perfectly competitive, firms producing a capital/consumption good by means of a constant returns to scale, time invariant technology, utilizing physical as well as human capital. The aggregate production function is given by

$$Y_t = AK_t^\alpha H_t^{1-\alpha}$$

(3)

where $A > 0$, $0 < \alpha < 1$, while $K_t$ is the aggregate supply of physical capital in period $t$. The coefficient $A$ characterizes the total factor productivity.

Perfect competition in the factors markets ensures that each factor is paid its marginal product so that the wage rate, $w_t$, and the rental rate, $R_t$, in period $t$, respectively, are given by

$$w_t = (1-\alpha)Y_t / H_t = (1-\alpha)AK_t^\alpha H_t^{-\alpha}$$

(4)

$$R_t = \alpha Y_t / K_t = \alpha AK_t^{\alpha-1} H_t^{1-\alpha}$$

(5)

2.3. Longevity

Individuals live for three periods. The first and second periods both have length 1. The third period length is individually determined. For individual $\omega$, the third period length is
\[ T(h_1(\omega)) \leq 1 \] thus, individual \( \omega \)’s longevity is \( 2 + T(h_1(\omega)) \). \( T(h_1(\omega)) \) is assumed to be an increasing function, so longevity increases with human capital. The function encompasses two assumptions: 1) human capital is defined as a single variable incorporating both investments in health and education, and 2) human capital investments are made during the period labeled ‘youth’.

The first assumption is justified by the fact that health and education are positively correlated, as observed in the data: higher levels of education are correlated with longer life expectancy. In fact, untangling the direction of causality, if any, is quite complex.\(^{24}\) The empirical evidence for a positive correlation between education and longevity is quite robust. Lillard and Waite (1995), Deaton and Paxson (1999), and Smith (1999) all find strong positive effects of education on life expectancy for the U.S. Similar findings are obtained by Attanasio and Emmerson (2001) for the U.K., von Gaudecker and Scholz (2007) for Germany, and Mesrine (1999) for France, among many others. Several causes for this correlation have been identified, some of which are causal (see Winkleby et al. (1992), Dewar (1998), Margettsa et al. (1998), Farbar and Levy (2000), and Lamerz et al. (2005)). The main causes identified are that higher educated individuals have better access to medical insurance, they are more able to utilize medical services, they eat healthier, and they have less physically demanding jobs. These causes have been identified to affect both individuals’ health and the health of their children.\(^{25}\)

Researchers continue to study why health and education are positively correlated. Recent research using general equilibrium models suggest that investments in health and education are complementary (see Blackburn and Cipriani (2002) and Finlay (2006)). Without using a causality assumption, this new research suggests that those who invest in

\(^{24}\) Grossman (2003) finds that education effects are causal but is less conclusive with regard to the identification of specific mechanisms. Fuchs (2004) discusses the complexities of disentangling the health/education relationship.

\(^{25}\) The relationship between health and education has been studied extensively over the last three decades with two notable works of Grossman (1972) and Fuchs (1982). Recent work by Bhattacharya and Qiao (2006), Chakraborty and Das (2005), Cutler and Lleras-Muney (2006), Kalemlı-Ozcan et al, (2000), Sanso and Aisa (2006), Zhang, Zhang, and Lee (2003), among many others, continue to provide evidence for such relationship.
education necessarily invest in health as well. Exploiting this complementary relationship, I permit human capital to incorporate both health and education investments.

The second assumption states that investments in human capital occur during youth. This is justifiable since the main emphasis of this paper is not on lifecycle healthcare behaviors, but on the distributional implications of unequal access to human capital investment. The youth period seems the most reasonable period for human capital investments as there is a clear pattern of investing in education during primary years, and empirical evidence indicates a strong causal effect of parents’ controlled input on children’s education and health (see Case, Lubotsky and Paxson, (2002), Case, Fertig, and Paxson (2005), Chen and Li (2009), and Desai and Alva (1998), among many others).

The assumption that longevity is a function of human capital or earnings capability continues to be used in the social security literature. In Bethencourt and Galasso (2001), the longevity function takes individuals’ ability type and public health as inputs to study the impact of Medicare on the size of Social Security.²⁶ Borck (2007) assumes longevity to be solely a function of one’s income to study different voting outcomes on social security taxes. Hachon (2008) uses a similar framework to Borck (2007) to analyze the progressivity of various social security systems.

2.4. Government

The government provides a social security system. This is financed through wage income tax revenue whereas the government must clear its budget each period. The social security system is annuity-based and PAYG, which redistributes revenue from current workers to current retirees. Contributions into the pension system are made through a constant and uniform tax, \( \tau \), on current workers’ earnings. During retirement, individuals receive a pension per unit of time until death. The pension individual \( \omega \) of generation \( t-1 \) receives during retirement is given by

²⁶ This is an extension of Philipson and Becker (1998) who analyze the complementarity of Medicare and Social Security. They find that individuals live ‘excessively’ due to Medicare providing a vehicle for individuals to live longer so as to gain social security benefits for a longer time.
where \( \nu_t \) is the benefit replacement rate in period \( t \) which is set to balance the government’s pension budget.

Given (6), the budget for the government’s pension system is

\[
\int_{\Omega_{t-1}} P_t(\omega) d\mu_{t-1}(\omega) = \tau w_t H_t
\]

(7)

2.5. Preferences and Budget Constraints

All individuals of generation \( t \) have identical intertemporal preferences over consumption when young, \( c_{t,t} \), consumption intensity (or per unit of time) when old, \( c_{t,t+1} \), and the level of human capital attainment of his child, \( h_{t+1} \), expressed by

\[
\ln c_{t,t} + T(h_t) \ln c_{t,t+1} + \delta \ln h_{t+1}
\]

(8)

where \( \delta \) is the degree of altruistic concern over the human capital of his offspring.\(^{28}\) As discussed in Section 2.3., \( T(h_t) \) is the length of the retirement period.

The budget constraints of individual \( \omega \) of generation \( t \) are

\[
c_{t,t}(\omega) + s_t(\omega) + e_t(\omega) = (1 - \tau)w_t h_t(\omega)
\]

(9)

\[
T(h_t(\omega)) c_{t,t+1}(\omega) = R_{t+1} s_t(\omega) + P_{t+1}(\omega)
\]

(10)

where \( w_t h_t(\omega) \) is the individual’s gross wage income, \( s_t(\omega) \) and \( e_t(\omega) \) are savings and private investment in children’s human capital, respectively.

2.6. Individual’s Maximization Problem

According to the model’s description, each individual \( \omega \) solves the problem

\[
\max \ln c_{t,t}(\omega) + T(h_t(\omega)) \ln c_{t,t+1}(\omega) + \delta \ln h_{t+1}(\omega)
\]

Similar to Zhang, Zhang, and Leung (2006), individuals equalize consumption at each point of time during working age and retirement leading to

\[
\int_0^T u(c_{t,t}) dt = u(c_{t,t}) \quad \text{and} \quad \int_0^T u(c_{t,t+1}) dt + 1 = Tu(c_{t,t+1})
\]

Note that equation (8) contains a warm glow formulation of imperfect altruism. See Osang and Sarkar (2007) for a discussion on the justification for this type of preferences.
\[ s.t. \quad c_{t,t} (\omega) + s_{t} (\omega) + e_{t} (\omega) = (1 - \tau) w_{t} h_{t} (\omega) \]
\[ T(h_{t} (\omega)) c_{t,t+1} (\omega) = R_{t+1} s_{t} (\omega) + P_{t+1} (\omega) \quad (11) \]
\[ h_{t+1} (\omega) = B e_{t} (\omega) \]

taking government policy variables, \( \tau, E_{t}, \) and \( P_{t+1} (\omega) \), and the values \( w_{t} \) and \( R_{t+1} \) as given. There is no nonnegativity constraint on \( s_{t} \).\(^{29}\) I now proceed to the general equilibrium analysis.

3. Definition of Dynamic Competitive Equilibrium

Given the human capital endowments, \( h_{0} (\omega) \) for all \( \omega \in \Omega_{0} \) and the initial physical capital stock, \( K_{-1} \), a dynamic competitive equilibrium is a collection of sequences of distributions of individual household decisions \( \{c_{t,t} (\omega), c_{t,t+1} (\omega), s_{t} (\omega), e_{t} (\omega)\}_{t=0}^{\infty} \), the sequences of aggregate amounts of physical capital and effective labor \( \{K_{t}, H_{t}\}_{t=0}^{\infty} \), the sequences of factor prices \( \{w_{t}, R_{t}\}_{t=0}^{\infty} \), and the sequences of distributions of government transfers \( \{P_{t} (\omega)\}_{t=0}^{\infty} \) such that:

1. **Individuals Solve Maximization Problem:** For each \( \omega \in \Omega_{t} \) and \( t = 0,1,..., \) the collection \( c_{t,t} (\omega), c_{t,t+1} (\omega), s_{t} (\omega), e_{t} (\omega), h_{t+1} (\omega) \) solves the individual’s problem (11)
2. **Competitive Factor Markets:** Factor prices are given by (4) and (5).
3. **Goods, Human Capital, and Physical Capital Markets Clear:** Markets clear so that the goods and aggregate stocks are given by

\[ Y_{t} = \int_{\Omega_{t}} c_{t,t} (\omega) d \mu_{t} (\omega) + \int_{\Omega_{t-1}} c_{t-1,t} (\omega) d \mu_{t-1} (\omega) \quad (12) \]
\[ H_{t} = \int_{\Omega_{t}} h_{t} (\omega) d \mu_{t} (\omega) \quad (13) \]
\[ K_{t} = \int_{\Omega_{t-1}} s_{t-1} (\omega) d \mu_{t-1} (\omega) \quad (14) \]

\(^{29}\) This is a simplification in the model since individuals know their social security pensions during their working years and there is no uncertainty about longevity of a specific individual.
(4) **Government Service:** Government pension budget is given by

\[ \int_{\Omega_{t-1}} P_t(\omega) d\mu_{t-1}(\omega) = \tau(1-\alpha)Y_t \]  

(15)

4. Solving the Model

Given the basic set-up and definition of dynamic general equilibrium, I now solve the model and provide the decision rules. The first order conditions for the individual’s maximization problem are

\[ \frac{1}{(1-\tau)w_t h_t(\omega) - e_t(\omega) - s_t(\omega)} = \frac{R_{t+1}}{R_{t+1}s_t(\omega) + P_{t+1}(\omega) - T(h_t(\omega))c_{t+1}(\omega)} \]  

(16)

\[ \frac{1}{(1-\tau)w_t h_t(\omega) - e_t(\omega) - s_t(\omega)} = \delta \]  

(17)

Using (16) and (17) along with (9) and (10) provides the decision rules:

\[ s_t(\omega) = \gamma_t(\omega)T(h_t(\omega))(1-\tau)w_t h_t(\omega) - \frac{\gamma_t(\omega)(1+\delta)P_{t+1}(\omega)}{R_{t+1}} \]  

(18)

and

\[ e_t(\omega) = \gamma_t(\omega)\delta \left( (1-\tau)w_t h_t(\omega) + \frac{P_{t+1}(\omega)}{R_{t+1}} \right) \]  

(19)

where \( \gamma_t(\omega) = (1+T(h_t(\omega))+\delta)^{-1} \). When longevity is independent of human capital attainment, all individuals have identical longevities, and thereby \( \gamma_t(\omega) = \gamma \). This implies that everyone spends identical fixed fractions on savings and investments in children’s human capital. When longevity is dependent on human capital attainment, \( \gamma \) is dependent on the length of retirement. It is easy to see that if the retirement period is short (long), then \( \gamma \) is large (small) and thereby, individuals invest more (less) on their children’s human capital and save less (more).\(^{30}\)

\(^{30}\)This dynamic is quite similar to Ehrlich and Lui (1991) who show that exogenous increases in longevity increase long-run growth.
To solve the model, I combine (7) with (4) and (5) to express the discounted social security pension in terms of savings

\[
\frac{P_{t+1}(\omega)}{R_{t+1}} = \frac{\tau_{t+1}w_{t+1}H_{t+1}}{\int_{\Omega} T(h_i(\omega))d\mu_i(\omega)}
\]

(20)

where

\[
\frac{w_{t+1}H_{t+1}}{R_{t+1}} = (1 - \alpha)\alpha^{-1}K_{t+1} = (1 - \alpha)\alpha^{-1}\int_{\Omega} s_i(\omega)d\mu(\omega)
\]

The first term on the right hand side of equation (20) is the inter-generational transfer which is equal to the present value of the aggregate social security benefits received by generation \( t \) retirees. The second term on the right hand side characterizes the degree of intra-generational redistribution among retirees by generation \( t \) retirees. When longevity is identical across all individuals, this term equals 1 which implies that social security benefits are uniformly distributed across retirees.

Substituting (20) into (18) yields individual \( \omega \)'s savings as a function of predetermined variables, exogenous parameters, and current period aggregate savings

\[
s_i(\omega) = T(h_i(\omega))\gamma_i(\omega)(1 - \tau)w_i(\omega) - \frac{(1 + \delta)\gamma,_i(\omega)T(h_i(\omega))\tau(1 - \alpha)}{\alpha\int_{\Omega} T(h_i(\omega))d\mu_i(\omega)}\int_{\Omega} s_i(\omega)d\mu(\omega)
\]

(21)

Integrating (21) over all individuals of generation \( t \) yields the equation for aggregate savings and next period’s physical capital stock in accordance to (14). The integration of (21) is analytically tractable when longevity is independent of human capital attainment and analytically untractable when longevity is dependent on human capital attainment. I first solve the model analytically, assuming that longevity is independent of human capital attainment and then assume that longevity is dependent on human capital attainment and proceed to numerical integrate (21) to solve for aggregate savings in order to study the effects of social security on welfare and growth.\(^{31}\)

Assuming that longevity is independent of human capital, the next period’s physical capital stock is

\[^{31}\text{I follow the solution method of Glomm and Kaganovich (2008) when longevity is independent of human capital attainment.}\]
Now using (20) in (1) with (19) and (22), I obtain the expression for the human capital of the children:

\[ h_{t+1}(\omega) = B\delta\gamma(1-\tau)(1-\alpha) \left[ \frac{h_i(\omega)}{H_i} + \psi \right] Y_i \]  

(23)

where

\[ \psi = \frac{T\tau(1-\alpha)\gamma}{\alpha + \tau(1-\alpha)(1-T\gamma)} \]

Integrating (23) over all individuals yields next period’s human capital stock

\[ H_{t+1} = B\delta\gamma(1-\tau)(1-\alpha) \left[ 1 + \psi \right] Y_i \]  

(24)

Using (22) and (24) in (3), I obtain the expression for stationary growth:

\[ g = \frac{D(1-\tau)C^{1-\alpha}}{T\alpha + (1+\delta)C} \]  

(25)

where \( C = \tau(1-\alpha) + \alpha \) and \( D = A(1-\alpha)(\delta B)^{1-\alpha}(\alpha T)^{\alpha} \).

Expression (25) makes it clear that, for certain parameter values, the size of the social security system will positively or negatively affect growth. The following condition assures that the growth rate is positive when social security is marginally increased from \( \tau = 0 \):\(^{32}\)

\[ T > \frac{(1+T+\delta+(1-\alpha)(1+\delta))\alpha}{(1-\alpha)^2} \]  

(26)

This condition makes it clear that if the concern for consumption in retirement, as indicated by the lengthen of retirement is sufficiently high, then social security enhances growth. The concern on consumption in retirement can be lower if the altruistic concern for children’s human capital or the return on physical capital is sufficiently low.

I now assume that longevity is dependent on human capital attainment. To solve the model numerically, I assume the following explicit deterministic functional form of the relationship between longevity and human capital attainment:

\[^{32}\text{This condition is invalid when } \tau \text{ is large.}\]
\[ T(h_t(\omega)) = \frac{ah_t(\omega)}{b + h_t(\omega)} \]  

where \( 0 \leq a \leq 1 \) and \( 0 \leq b < \infty \). The parameter \( a \) determines the upper bound on the length of the retirement period. The parameter \( b \) affects the dependency of longevity on human capital. When \( b = 0 \), longevity and human capital are independent and all individuals have identical longevities equal to length \( a \). This corresponds to the commonly used constant probability of survival function in the literature. When \( b > 0 \), longevity is dependent on human capital and has the following properties: \( T'(h) > 0 \), \( T''(h) < 0 \), \( T(0) = 0 \), \( \lim_{h \to \infty} T(h) \to a \). The function is concave which permits diminishing returns for longevity from increases in human capital. This is in line with the data which reveal that increases in education and health investments from lower levels tend to produce relatively larger increases in longevity. gasoline. Lastly, when \( a = b = 1 \), (27) is the probability of survival function as in Blackburn and Cipriani (2002), Chakraborty (2004), Finlay (2006), and Osang and Sarkar (2007).

In the next section present the results to the numerical analysis of the model. Rather than pursuing a calibration exercise, I set parameter values which are consistent in the findings in the literature. gasoline. I set the total factor productivity parameters for goods production and human capital production equal to 3.8. This is close to the minimum values required for growth to be sustained in all simulations. For the preference parameter, I assume an altruism value of \( \delta = .5 \). This value forces individuals to care more about consumption than their child’s human capital. The initial human capital distribution is assumed to be log-normally distributed with a mean of 1 and a variance of 1. The longevity parameter is set at \( a = .6 \) corresponding to U.S. life expectancy data for 2005-2015. (See Table A.17 of the United Nations World Population Prospects: 2006

---


34 There is one parameter which is non-standard from the literature, \( \delta \). The altruism parameter, \( \delta \), does not have a reliable estimate.
The value of \( b \) is taken to be set at either 0 corresponding with longevity being independent of human capital attainment or a value of 1 corresponding to longevity being dependent on human capital attainment. This permits sufficient dispersion in longevity. Lastly, I set the capital share to assure condition (26) holds.

In all simulations, I numerically verify that the equilibrium value of aggregate savings is uniquely defined in all periods and that all markets clear. I compute the growth rate as the percentage change in output from one period to the next and measure all inequalities using the Gini coefficient.

5. Analysis of the Model

In this section I conduct policy experiments when a PAYG Beveridgian type social security scheme is introduced. I first assume that longevity is independent of human capital and study the effects of implementing social security on welfare inequality and growth. I then assume that longevity is dependent on human capital attainment and study the effects of implementing social security on welfare inequality and growth.

**Numerical Result 5.1.** Assume that the parametric assumptions stated in Section 4 hold and that longevity is independent of human capital attainment, i.e., it is identical across all individuals. Then if a Beveridgian social security system is implemented at time \( t = 0 \), there exists a threshold tax rate \( \hat{\tau} \in (0, 1) \) such that the following is true in comparison to the baseline case without social security:

(i) If the social security tax, \( \tau \), is less than or equal to \( \hat{\tau} \), then aggregate human capital and output are higher in all subsequent periods \( t = 1, 2, K \).

(ii) If the social security tax, \( \tau \), is greater than or equal to \( \tau \), then aggregate human capital is higher while aggregate physical capital and output are lower in all subsequent periods \( t = 1, 2, K \).

---

35 One period duration is 30 years so that people retire at age 60. When longevity is independent of human capital, the average life expectancy for the initial working-age generation is 78 years.
Furthermore, for all \( \tau \), distributions of welfare, human capital attainment, and longevity become strictly more equal in all subsequent periods \( t = 1, 2, K \).

This is the benchmark result which is known in the literature (see Kaganovich and Zilcha (1999) and Glomm and Kaganovich (2008). If all individuals have identical longevity, which do not change over time, then propensities to invest in children’s human capital and to save are also identical across individuals and constant over time. The result of implementing a Beveridgian social security system is that the inequalities in human capital attainment and welfare become strictly more equal while growth may increase or decrease compared to an economy without social security. The Beveridgian formula provides additional resources to those with lower levels of human capital. These resources are used to increase consumption, savings, and investments in their children’s human capital. Thus, implementing social security leads to a more equal distribution of human capital and welfare.

The effect of social security on growth depends critically on the size of social security system. If the social security system of small magnitude is introduced, this will enhance growth. The main channel to higher growth is through increasing the stock of human capital. The PAYG design creates incentives for current workers to invest in the next generation’s human capital, to increase their own social security benefits in retirement. If the increase in children’s human capital is sufficiently large, it compensates for the decrease in savings caused by the PAYG design, and output is enhanced. When the social security system is large, the increase in human capital is not sufficient to overcome the loss of savings. This causes output to decline. I now proceed to analyze the comparative dynamics when longevity is dependent on human capital attainment.

The result that social security enhances growth is conditional upon the chosen parameter value given in Section 4. Condition (26) must be satisfied, in other words, individuals’ concern for consumption in retirement must be sufficiently strong. Similar conditions are provided in Kaganovich and Zilcha (1999) and Glomm and Kaganovich (2008) when financing social security competes with financing public education. My computational results indicate that, for the chosen parameter values, the threshold, \( \hat{\tau} \), is
approximately 16%. The tax threshold, \( \hat{\tau} \), adjusts downward when the degree of altruism or the return on physical capital are larger. It should be quite clear that if condition (26) is not satisfied, then social security will not enhance growth regardless of the size of social security. Since this paper analyzes the effect of the dependence assumption of longevity on human capital attainment, I assume condition (26) holds for all experiments.

**Numerical Result 5.2.** Assume that the parametric assumptions in Section 4 hold and that longevity depends on human capital attainment according to expression (27). Then if a Beveridgian social security system is implemented at time \( t = 0 \), there exist threshold tax rates \( \underline{\tau}, \bar{\tau} \in (0, 1) \) where \( \underline{\tau} < \bar{\tau} \) and a finite period \( s \) such that, the following is true in comparison to the baseline case without social security:

(i) If the social security tax, \( \tau \), is less than \( \underline{\tau} \), then aggregate human capital and output are higher in all subsequent periods \( t = 1, 2, K \)

(ii) If the social security tax is \( \tau \) is between \( \underline{\tau} \) and \( \bar{\tau} \), then
   a. aggregate human capital is higher while output is lower in periods \( t = 1, 2, K, s \)
   b. aggregate human capital and output are higher while physical capital is lower in all subsequent periods \( t = s + 1, s + 2, K \)

(iii) If the social security tax, \( \tau \), is greater than \( \bar{\tau} \), then aggregate human capital is higher while physical capital and output are lower in all subsequent periods \( t = 1, 2, K \);

Furthermore, for all \( \tau \), distributions of welfare, human capital attainment, and longevity become strictly more equal in all subsequent periods \( t = 1, 2, K \);

Similar to Numerical Result 5.1., if the social security system is sufficiently small, it enhances growth. When social security is implemented, investments in human capital of the younger generation rise in response to the new incentives to invest in children’s human capital for higher social security benefits. Individuals decrease their savings in
response to receiving social security benefits during retirement. Since the increase in human capital is greater than the loss of savings, output rises.

Contrary to the benchmark results, as the social security system becomes larger, the results exhibit an intertemporal non-monotonic relationship between growth and welfare inequality. In the short-run, welfare inequality decreases along with output while, in the long-run, output increases. There are two effects from implementing social security which drive this result, one which is standard and the other which is a new insight provided by this paper.

The first one is that the PAYG Beveridgian formula causes individuals to increase investments in the younger generation’s human capital and save less for retirement. The Beveridgian formula causes lifetime resources of the poor to increase while the lifetime resources of wealthy decrease. This, in turn, increases investment in children’s human capital of the poor and decreases investment in children’s human capital of the wealthy.\(^{36}\) Since poorer individuals have higher propensities to invest in their children’s human capital, the increase in investment in children’s human capital of the poor is greater than the decrease in investment of the wealthy. Thus, aggregate human capital of the younger generation rises while the distribution of human capital becomes more equal. When social security is large, the increase in human capital is not sufficient to compensate for the decrease in savings from the PAYG design, lowering output in the short-run.

The second effect is the dynamic externality of lengthened children’s longevity due to increasing their human capital. As longevity lengthens, individuals increase the fraction of resources devoted to savings while reducing the fraction devoted to investment in children’s human capital. The overall effect on human capital accumulation and savings is conditional on how much longevity lengthens. To provide intuition on these results, I update equations (18) and (19) one period and partially differentiate them with respect to human capital. This provides the change in savings and investments in

\(^{36}\) The specification of the human capital production function can significantly change this result. For instance, in Glomm and Kaganovich (2008) private investments in human capital are in the form of a time input. Implementing social security results in all individuals increasing their time input with their children.
children’s human capital when the younger generation’s human capital is increased. Differentiating (18) and (19) with respect to $h_t$ yields:

\[
\frac{\partial s_{t+1}(\omega)}{\partial h_{t+1}(\omega)} = \gamma_{t+1}(\omega)T(h_{t+1}(\omega))(1-\tau)w_{t+1}
\]

\[
-\frac{(1+\delta)\gamma_{t+1}(\omega)T(h_{t+1}(\omega))}{R_{t+2}} \left( \frac{\partial P_{t+2}(\omega)}{\partial h_{t+1}(\omega)} \frac{\partial T(h_{t+1}(\omega))}{\partial h_{t+1}(\omega)} \right) + (1+\delta)\gamma_{t+1}(\omega)^2 \left( (1-\tau)w_{t+1}h_{t+1}(\omega) - \left( \frac{P_{t+2}(\omega)}{R_{t+2}} \right) \right) \frac{\partial T(h_{t+1}(\omega))}{\partial h_{t+1}(\omega)}
\]

and

\[
\frac{\partial e_{t+1}(\omega)}{\partial h_{t+1}(\omega)} = \gamma_{t+1}(\omega)\delta(1-\tau)w_{t+1} + \gamma_{t+1}(\omega)\delta \left( \frac{\partial P_{t+2}(\omega)}{\partial h_{t+1}(\omega)} \right)
\]

\[
-\delta\gamma_{t+1}(\omega)^2 \left( (1-\tau)w_{t+1}h_{t+1}(\omega) + \left( \frac{P_{t+2}(\omega)}{R_{t+2}} \right) \right) \frac{\partial T(h_{t+1}(\omega))}{\partial h_{t+1}(\omega)}
\]

The first term in (28) is the direct effect of increasing children’s human capital on their savings. As children’s human capital rises, concern for their retirement increases causing their savings to rise. The second term in (28) is the effect of increasing social security benefits on savings. This is standard in the literature, as social security benefits rise (in this case due to a longer retirement period) individuals save less. The third term in (28) is the effect of the dynamic externality on savings. As longevity increases, individuals save more for retirement.

Similarly, the first term in (29) is the direct effect of increasing children’s human capital on investment in their children’s human capital. As children’s human capital rises, additional income is available to increase investment in their children’s human capital. The second term in (29) is the effect of increasing social security benefits on investments in their children’s human capital. As pension benefits increase, lifetime resources increase and, in part, are used to increase investment in human capital. The third term in (29) is the effect of the dynamic externality. As longevity increases, it causes individuals to invest less in their children’s human capital and more toward their retirement.

There are a few comments which need to be discussed. The first is that both effects are much stronger with respect to savings than investments in human capital. The
second is that diminishing returns on longevity from increasing human capital imply that, at some point, the second and third term of both (28) and (29) will approach zero. The third is that the concavity of the longevity function in human capital attainment implies that the effect of social security and the dynamic externality are larger for those with lower levels of human capital than those with higher levels of human capital. Fourth, if the increase in human capital is sufficiently large, then the effect of social security and the dynamic externality on investments in human capital will not decrease human capital of the next generation. Even if this holds for investments in human capital, this does not have to hold for savings. Thus, in the long-run, implementing PAYG social security can cause the stock of human capital to grow while savings to increase due to the dynamic externality, raising output in the long-run.

When social security becomes excessively large, welfare inequality declines at a cost of lower growth in the short-run and in the long-run. If the tax rate becomes too high, then the increase in children’s human capital is not sufficient to compensate for the large reduction in savings. Even with future generations’ increased concern for retirement, the stock of human capital and savings do not rise sufficiently to compensate for the loss of savings from the large social security in production.

The chosen parameter values provide the period and tax thresholds in which social security enhances growth. The tax threshold, $\tau$, which guarantees that social security enhances growth is 14% while the other tax threshold, $\bar{\tau}$, is 17%. For a tax rate of 15%, growth will be enhanced after the second period social security is implemented while after the fifth period output is higher. Assuming that a period length is 30 years, the positive effects on growth of implementing social security take roughly 60 years to be effective. If the model had more working periods, this would shorten the number of periods required before growth is enhanced.

6. Conclusion
In this paper, I explore the effects of implementing a PAYG Beveridgian social security system on growth and welfare inequality, when longevity may depend on human capital
attainment. To study these dynamical effects, I construct a model of overlapping generations populated by heterogeneous individuals. I assume individuals are heterogeneous due to their parents’ investments in their human capital. When longevity is independent on human capital attainment, standard benchmark results are established, which are also found in the literature. I then extend the framework to permit longevity to be dependent on human capital attainment. When longevity is dependent on human capital, implementing a PAYG social security system reduces inequality but additionally raises the longevity of the next generation through increases in human capital attainment. This inter-temporal externality has implications on the propensities to save and invest in human capital.

The analysis reveals that the assumption on the dependency of longevity on human capital attainment and the size of social security play dominant roles in determining the effect on growth. When the social security system is small, implementing social security enhances growth regardless of the dependency assumption of longevity on human capital attainment. As the social security system becomes large, the decline in savings becomes increasingly more difficult to compensate in production. The dynamic externality of raising longevity, due to increasing human capital, raises individuals concern for retirement and causing them to save more. This compensates for the negative effect of PAYG social security on growth and results in a temporally non-monotonic growth path: in the short-run, growth declines while in the long-run, growth is higher. For excessively large social security system, growth declines along the entire transition path regardless of the dependency assumption of longevity on human capital attainment.

While this model makes a step toward a better understanding of the impact of endogenous longevity on growth and welfare inequality, further research is needed to extend the endogeneity of individuals’ longevity. Models which permit individuals to directly influence their own health through preventative medical services during adulthood and life-extending medical procedures during retirement are required to more fully capture the effects of social security.
References


Appendix A: Numerical Result 5.1.

Table A.1.: Numerical Result 5.1.

<table>
<thead>
<tr>
<th>$t$</th>
<th>$\tau$</th>
<th>$\text{Gini}(h)$</th>
<th>$\text{Gini}(T)$</th>
<th>$\text{Gini}(U)$</th>
<th>$T$</th>
<th>$H$</th>
<th>$K$</th>
<th>$Y$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>0.00</td>
<td>51.6</td>
<td>61.7</td>
<td>2.6</td>
<td>444</td>
<td>144</td>
<td>1,510</td>
<td></td>
</tr>
<tr>
<td>1.00</td>
<td>0.00</td>
<td>51.6</td>
<td>41.9</td>
<td>2.6</td>
<td>619</td>
<td>543</td>
<td>2,320</td>
<td></td>
</tr>
<tr>
<td>2.00</td>
<td>0.00</td>
<td>51.6</td>
<td>31.7</td>
<td>2.6</td>
<td>952</td>
<td>835</td>
<td>3,570</td>
<td></td>
</tr>
<tr>
<td>3.00</td>
<td>0.00</td>
<td>51.6</td>
<td>25.5</td>
<td>2.6</td>
<td>1,466</td>
<td>1,286</td>
<td>5,500</td>
<td></td>
</tr>
<tr>
<td>4.00</td>
<td>0.00</td>
<td>51.6</td>
<td>21.3</td>
<td>2.6</td>
<td>2,256</td>
<td>1,979</td>
<td>8,460</td>
<td></td>
</tr>
<tr>
<td>5.00</td>
<td>0.00</td>
<td>51.6</td>
<td>18.3</td>
<td>2.6</td>
<td>3,473</td>
<td>3,046</td>
<td>13,020</td>
<td></td>
</tr>
<tr>
<td>6.00</td>
<td>0.00</td>
<td>51.6</td>
<td>16.1</td>
<td>2.6</td>
<td>5,345</td>
<td>4,689</td>
<td>20,050</td>
<td></td>
</tr>
<tr>
<td>7.00</td>
<td>0.00</td>
<td>51.6</td>
<td>14.3</td>
<td>2.6</td>
<td>8,227</td>
<td>7,217</td>
<td>30,860</td>
<td></td>
</tr>
<tr>
<td>8.00</td>
<td>0.00</td>
<td>51.6</td>
<td>12.9</td>
<td>2.6</td>
<td>12,664</td>
<td>11,109</td>
<td>47,500</td>
<td></td>
</tr>
<tr>
<td>9.00</td>
<td>0.00</td>
<td>51.6</td>
<td>11.7</td>
<td>2.6</td>
<td>19,492</td>
<td>17,098</td>
<td>73,110</td>
<td></td>
</tr>
<tr>
<td>10.00</td>
<td>0.00</td>
<td>51.6</td>
<td>10.8</td>
<td>2.6</td>
<td>30,033</td>
<td>26,318</td>
<td>112,530</td>
<td></td>
</tr>
<tr>
<td>11.00</td>
<td>0.00</td>
<td>51.6</td>
<td>9.9</td>
<td>2.6</td>
<td>46,180</td>
<td>40,510</td>
<td>173,200</td>
<td></td>
</tr>
<tr>
<td>12.00</td>
<td>0.00</td>
<td>51.6</td>
<td>9.2</td>
<td>2.6</td>
<td>71,080</td>
<td>62,350</td>
<td>266,600</td>
<td></td>
</tr>
<tr>
<td>13.00</td>
<td>0.00</td>
<td>51.6</td>
<td>8.6</td>
<td>2.6</td>
<td>109,410</td>
<td>95,980</td>
<td>410,400</td>
<td></td>
</tr>
<tr>
<td>14.00</td>
<td>0.00</td>
<td>51.6</td>
<td>8.1</td>
<td>2.6</td>
<td>168,410</td>
<td>147,730</td>
<td>631,600</td>
<td></td>
</tr>
<tr>
<td>15.00</td>
<td>0.00</td>
<td>51.6</td>
<td>7.6</td>
<td>2.6</td>
<td>259,220</td>
<td>227,380</td>
<td>972,200</td>
<td></td>
</tr>
</tbody>
</table>

Table A.1.: Baseline Transition Path Under The Assumption of Identical Longevity
(No Social Security)

Numerical Result 5.1.i.: Transition Path With Beveridgian Social Security Under The Assumption of Identical Longevity

<table>
<thead>
<tr>
<th>$t$</th>
<th>$\tau$</th>
<th>$\text{Gini}(h)$</th>
<th>$\text{Gini}(T)$</th>
<th>$\text{Gini}(U)$</th>
<th>$T$</th>
<th>$H$</th>
<th>$K$</th>
<th>$Y$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>0.00</td>
<td>51.6</td>
<td>51.8</td>
<td>2.6</td>
<td>444</td>
<td>144</td>
<td>1,510</td>
<td></td>
</tr>
<tr>
<td>1.05</td>
<td>0.00</td>
<td>45.2</td>
<td>25.8</td>
<td>2.6</td>
<td>671</td>
<td>406</td>
<td>2,420</td>
<td></td>
</tr>
<tr>
<td>2.05</td>
<td>0.00</td>
<td>39.5</td>
<td>15.8</td>
<td>2.6</td>
<td>1,079</td>
<td>653</td>
<td>3,900</td>
<td></td>
</tr>
<tr>
<td>3.05</td>
<td>0.00</td>
<td>34.6</td>
<td>10.8</td>
<td>2.6</td>
<td>1,736</td>
<td>1,050</td>
<td>6,270</td>
<td></td>
</tr>
<tr>
<td>4.05</td>
<td>0.00</td>
<td>30.3</td>
<td>7.8</td>
<td>2.6</td>
<td>2,793</td>
<td>1,690</td>
<td>10,090</td>
<td></td>
</tr>
<tr>
<td>5.05</td>
<td>0.00</td>
<td>26.6</td>
<td>5.8</td>
<td>2.6</td>
<td>4,493</td>
<td>2,718</td>
<td>16,240</td>
<td></td>
</tr>
<tr>
<td>6.05</td>
<td>0.00</td>
<td>23.3</td>
<td>4.4</td>
<td>2.6</td>
<td>7,228</td>
<td>4,373</td>
<td>26,120</td>
<td></td>
</tr>
<tr>
<td>7.05</td>
<td>0.00</td>
<td>20.4</td>
<td>3.5</td>
<td>2.6</td>
<td>11,627</td>
<td>7,034</td>
<td>42,020</td>
<td></td>
</tr>
<tr>
<td>8.05</td>
<td>0.00</td>
<td>17.9</td>
<td>2.8</td>
<td>2.6</td>
<td>18,703</td>
<td>11,314</td>
<td>67,590</td>
<td></td>
</tr>
<tr>
<td>9.05</td>
<td>0.00</td>
<td>15.6</td>
<td>2.2</td>
<td>2.6</td>
<td>30,086</td>
<td>18,201</td>
<td>108,720</td>
<td></td>
</tr>
<tr>
<td>10.05</td>
<td>0.00</td>
<td>13.7</td>
<td>1.8</td>
<td>2.6</td>
<td>48,396</td>
<td>29,278</td>
<td>174,890</td>
<td></td>
</tr>
<tr>
<td>11.05</td>
<td>0.00</td>
<td>12.0</td>
<td>1.5</td>
<td>2.6</td>
<td>77,900</td>
<td>47,100</td>
<td>281,300</td>
<td></td>
</tr>
<tr>
<td>12.05</td>
<td>0.00</td>
<td>10.5</td>
<td>1.2</td>
<td>2.6</td>
<td>125,200</td>
<td>75,760</td>
<td>452,600</td>
<td></td>
</tr>
<tr>
<td>13.05</td>
<td>0.00</td>
<td>9.2</td>
<td>1.0</td>
<td>2.6</td>
<td>201,400</td>
<td>121,870</td>
<td>728,000</td>
<td></td>
</tr>
<tr>
<td>14.05</td>
<td>0.00</td>
<td>8.1</td>
<td>0.8</td>
<td>2.6</td>
<td>324,100</td>
<td>196,040</td>
<td>1,171,000</td>
<td></td>
</tr>
<tr>
<td>15.05</td>
<td>0.00</td>
<td>7.1</td>
<td>0.7</td>
<td>2.6</td>
<td>521,300</td>
<td>315,350</td>
<td>1,883,800</td>
<td></td>
</tr>
</tbody>
</table>
### Numerical Result 5.1.i.: Baseline Transition Path Under The Assumption of Identical Longevity (No Social Security)

<table>
<thead>
<tr>
<th>t</th>
<th>τ</th>
<th>Gini(h)</th>
<th>Gini(T)</th>
<th>Gini(U)</th>
<th>T</th>
<th>H</th>
<th>K</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.10</td>
<td>51.6</td>
<td>0.0</td>
<td>38.2</td>
<td>2.6</td>
<td>444</td>
<td>144</td>
<td>1,510</td>
</tr>
<tr>
<td>1</td>
<td>0.10</td>
<td>41.8</td>
<td>0.0</td>
<td>18.8</td>
<td>2.6</td>
<td>687</td>
<td>317</td>
<td>2,420</td>
</tr>
<tr>
<td>2</td>
<td>0.10</td>
<td>33.9</td>
<td>0.0</td>
<td>11.2</td>
<td>2.6</td>
<td>1,101</td>
<td>508</td>
<td>3,870</td>
</tr>
<tr>
<td>3</td>
<td>0.10</td>
<td>27.5</td>
<td>0.0</td>
<td>7.3</td>
<td>2.6</td>
<td>1,765</td>
<td>815</td>
<td>6,210</td>
</tr>
<tr>
<td>4</td>
<td>0.10</td>
<td>22.3</td>
<td>0.0</td>
<td>5.0</td>
<td>2.6</td>
<td>2,829</td>
<td>1,306</td>
<td>9,950</td>
</tr>
<tr>
<td>5</td>
<td>0.10</td>
<td>18.0</td>
<td>0.0</td>
<td>3.5</td>
<td>2.6</td>
<td>4,535</td>
<td>2,094</td>
<td>15,950</td>
</tr>
<tr>
<td>6</td>
<td>0.10</td>
<td>14.6</td>
<td>0.0</td>
<td>2.6</td>
<td>2.6</td>
<td>7,268</td>
<td>3,356</td>
<td>25,570</td>
</tr>
<tr>
<td>7</td>
<td>0.10</td>
<td>11.8</td>
<td>0.0</td>
<td>1.9</td>
<td>2.6</td>
<td>11,650</td>
<td>5,379</td>
<td>40,980</td>
</tr>
<tr>
<td>8</td>
<td>0.10</td>
<td>9.6</td>
<td>0.0</td>
<td>1.4</td>
<td>2.6</td>
<td>18,674</td>
<td>8,621</td>
<td>65,680</td>
</tr>
<tr>
<td>9</td>
<td>0.10</td>
<td>7.8</td>
<td>0.0</td>
<td>1.0</td>
<td>2.6</td>
<td>29,932</td>
<td>13,819</td>
<td>105,280</td>
</tr>
<tr>
<td>10</td>
<td>0.10</td>
<td>6.3</td>
<td>0.0</td>
<td>0.8</td>
<td>2.6</td>
<td>47,978</td>
<td>22,150</td>
<td>168,750</td>
</tr>
<tr>
<td>11</td>
<td>0.10</td>
<td>5.1</td>
<td>0.0</td>
<td>0.6</td>
<td>2.6</td>
<td>76,900</td>
<td>35,500</td>
<td>270,500</td>
</tr>
<tr>
<td>12</td>
<td>0.10</td>
<td>4.1</td>
<td>0.0</td>
<td>0.5</td>
<td>2.6</td>
<td>123,300</td>
<td>56,910</td>
<td>433,600</td>
</tr>
<tr>
<td>13</td>
<td>0.10</td>
<td>3.4</td>
<td>0.0</td>
<td>0.4</td>
<td>2.6</td>
<td>197,600</td>
<td>91,220</td>
<td>695,000</td>
</tr>
<tr>
<td>14</td>
<td>0.10</td>
<td>2.7</td>
<td>0.0</td>
<td>0.3</td>
<td>2.6</td>
<td>316,700</td>
<td>146,210</td>
<td>1,113,900</td>
</tr>
<tr>
<td>15</td>
<td>0.10</td>
<td>2.2</td>
<td>0.0</td>
<td>0.2</td>
<td>2.6</td>
<td>507,600</td>
<td>234,360</td>
<td>1,785,500</td>
</tr>
</tbody>
</table>

### Numerical Result 5.1.ii.: Transition Path With Beveridgean Social Security Under The Assumption of Identical Longevity

<table>
<thead>
<tr>
<th>t</th>
<th>τ</th>
<th>Gini(h)</th>
<th>Gini(T)</th>
<th>Gini(U)</th>
<th>T</th>
<th>H</th>
<th>K</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.17</td>
<td>51.6</td>
<td>0.0</td>
<td>31.1</td>
<td>2.6</td>
<td>444</td>
<td>144</td>
<td>1,510</td>
</tr>
<tr>
<td>1</td>
<td>0.17</td>
<td>39.1</td>
<td>0.0</td>
<td>15.4</td>
<td>2.6</td>
<td>677</td>
<td>235</td>
<td>2,320</td>
</tr>
<tr>
<td>2</td>
<td>0.17</td>
<td>29.6</td>
<td>0.0</td>
<td>9.0</td>
<td>2.6</td>
<td>1,040</td>
<td>361</td>
<td>3,560</td>
</tr>
<tr>
<td>3</td>
<td>0.17</td>
<td>22.5</td>
<td>0.0</td>
<td>5.6</td>
<td>2.6</td>
<td>1,598</td>
<td>554</td>
<td>5,460</td>
</tr>
<tr>
<td>4</td>
<td>0.17</td>
<td>17.0</td>
<td>0.0</td>
<td>3.7</td>
<td>2.6</td>
<td>2,454</td>
<td>851</td>
<td>8,390</td>
</tr>
<tr>
<td>5</td>
<td>0.17</td>
<td>12.9</td>
<td>0.0</td>
<td>2.5</td>
<td>2.6</td>
<td>3,768</td>
<td>1,307</td>
<td>12,880</td>
</tr>
<tr>
<td>6</td>
<td>0.17</td>
<td>9.8</td>
<td>0.0</td>
<td>1.7</td>
<td>2.6</td>
<td>5,788</td>
<td>2,007</td>
<td>19,780</td>
</tr>
<tr>
<td>7</td>
<td>0.17</td>
<td>7.4</td>
<td>0.0</td>
<td>1.2</td>
<td>2.6</td>
<td>8,889</td>
<td>3,082</td>
<td>30,380</td>
</tr>
<tr>
<td>8</td>
<td>0.17</td>
<td>5.6</td>
<td>0.0</td>
<td>0.8</td>
<td>2.6</td>
<td>13,651</td>
<td>4,733</td>
<td>46,660</td>
</tr>
<tr>
<td>9</td>
<td>0.17</td>
<td>4.3</td>
<td>0.0</td>
<td>0.6</td>
<td>2.6</td>
<td>20,966</td>
<td>7,269</td>
<td>71,660</td>
</tr>
<tr>
<td>10</td>
<td>0.17</td>
<td>3.2</td>
<td>0.0</td>
<td>0.4</td>
<td>2.6</td>
<td>32,200</td>
<td>11,164</td>
<td>110,060</td>
</tr>
<tr>
<td>11</td>
<td>0.17</td>
<td>2.5</td>
<td>0.0</td>
<td>0.3</td>
<td>2.6</td>
<td>494,500</td>
<td>17,150</td>
<td>169,000</td>
</tr>
<tr>
<td>12</td>
<td>0.17</td>
<td>1.9</td>
<td>0.0</td>
<td>0.2</td>
<td>2.6</td>
<td>759,500</td>
<td>26,330</td>
<td>259,600</td>
</tr>
<tr>
<td>13</td>
<td>0.17</td>
<td>1.4</td>
<td>0.0</td>
<td>0.2</td>
<td>2.6</td>
<td>1,166,500</td>
<td>40,440</td>
<td>398,700</td>
</tr>
<tr>
<td>14</td>
<td>0.17</td>
<td>1.1</td>
<td>0.0</td>
<td>0.1</td>
<td>2.6</td>
<td>1,791,500</td>
<td>62,110</td>
<td>612,300</td>
</tr>
<tr>
<td>15</td>
<td>0.17</td>
<td>0.8</td>
<td>0.0</td>
<td>0.1</td>
<td>2.6</td>
<td>2,751,400</td>
<td>95,390</td>
<td>940,400</td>
</tr>
</tbody>
</table>
Table A.1. con.: Numerical Result 5.1.

Numerical Result 5.1.i.: Transition Path With Beveridgian Social Security Under The Assumption of Identical Longevity

<table>
<thead>
<tr>
<th>$t$</th>
<th>$\tau$</th>
<th>$\text{Gini}(h)$</th>
<th>$\text{Gini}(T)$</th>
<th>$\text{Gini}(U)$</th>
<th>$T$</th>
<th>$H$</th>
<th>$K$</th>
<th>$Y$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.18</td>
<td>51.6</td>
<td>0.0</td>
<td>30.5</td>
<td>2.6</td>
<td>444</td>
<td>1,440</td>
<td>1,510</td>
</tr>
<tr>
<td>1</td>
<td>0.18</td>
<td>38.8</td>
<td>0.0</td>
<td>15.1</td>
<td>2.6</td>
<td>674</td>
<td>2,257</td>
<td>2,300</td>
</tr>
<tr>
<td>2</td>
<td>0.18</td>
<td>29.2</td>
<td>0.0</td>
<td>8.8</td>
<td>2.6</td>
<td>1,026</td>
<td>3,437</td>
<td>3,500</td>
</tr>
<tr>
<td>3</td>
<td>0.18</td>
<td>22.0</td>
<td>0.0</td>
<td>5.5</td>
<td>2.6</td>
<td>1,563</td>
<td>5,234</td>
<td>5,320</td>
</tr>
<tr>
<td>4</td>
<td>0.18</td>
<td>16.5</td>
<td>0.0</td>
<td>3.6</td>
<td>2.6</td>
<td>2,381</td>
<td>7,971</td>
<td>8,110</td>
</tr>
<tr>
<td>5</td>
<td>0.18</td>
<td>12.5</td>
<td>0.0</td>
<td>2.4</td>
<td>2.6</td>
<td>3,626</td>
<td>12,139</td>
<td>12,350</td>
</tr>
<tr>
<td>6</td>
<td>0.18</td>
<td>9.4</td>
<td>0.0</td>
<td>1.6</td>
<td>2.6</td>
<td>5,522</td>
<td>18,486</td>
<td>18,810</td>
</tr>
<tr>
<td>7</td>
<td>0.18</td>
<td>7.1</td>
<td>0.0</td>
<td>1.1</td>
<td>2.6</td>
<td>8,409</td>
<td>28,153</td>
<td>28,640</td>
</tr>
<tr>
<td>8</td>
<td>0.18</td>
<td>5.3</td>
<td>0.0</td>
<td>0.8</td>
<td>2.6</td>
<td>12,806</td>
<td>42,876</td>
<td>43,620</td>
</tr>
<tr>
<td>9</td>
<td>0.18</td>
<td>4.0</td>
<td>0.0</td>
<td>0.6</td>
<td>2.6</td>
<td>19,503</td>
<td>65,296</td>
<td>66,430</td>
</tr>
<tr>
<td>10</td>
<td>0.18</td>
<td>3.0</td>
<td>0.0</td>
<td>0.4</td>
<td>2.6</td>
<td>29,701</td>
<td>99,441</td>
<td>101,170</td>
</tr>
<tr>
<td>11</td>
<td>0.18</td>
<td>2.3</td>
<td>0.0</td>
<td>0.3</td>
<td>2.6</td>
<td>45,230</td>
<td>154,100</td>
<td>154,100</td>
</tr>
<tr>
<td>12</td>
<td>0.18</td>
<td>1.7</td>
<td>0.0</td>
<td>0.2</td>
<td>2.6</td>
<td>688,900</td>
<td>23,060</td>
<td>234,620</td>
</tr>
<tr>
<td>13</td>
<td>0.18</td>
<td>1.3</td>
<td>0.0</td>
<td>0.1</td>
<td>2.6</td>
<td>1,049,100</td>
<td>35,120</td>
<td>357,300</td>
</tr>
<tr>
<td>14</td>
<td>0.18</td>
<td>1.0</td>
<td>0.0</td>
<td>0.1</td>
<td>2.6</td>
<td>1,597,700</td>
<td>53,490</td>
<td>544,200</td>
</tr>
<tr>
<td>15</td>
<td>0.18</td>
<td>0.7</td>
<td>0.0</td>
<td>0.1</td>
<td>2.6</td>
<td>2,433,100</td>
<td>81,460</td>
<td>828,800</td>
</tr>
</tbody>
</table>

125
Appendix B: Numerical Result 5.2.

### Table B.1.: Numerical Result 5.2.

**Numerical Result 5.2.: Baseline Transition Path Under The Assumption Longevity is Dependent on Human Capital Attainment (No Social Security)**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>Gini(h)</th>
<th>Gini(T)</th>
<th>Gini(U)</th>
<th>T</th>
<th>H</th>
<th>K</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>51.6</td>
<td>14.6</td>
<td>49.0</td>
<td>2.42</td>
<td>444</td>
<td>144</td>
<td>1,510</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>46.0</td>
<td>8.1</td>
<td>29.4</td>
<td>2.50</td>
<td>711</td>
<td>485</td>
<td>2,600</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>42.1</td>
<td>4.4</td>
<td>20.7</td>
<td>2.55</td>
<td>1,176</td>
<td>868</td>
<td>4,340</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>39.7</td>
<td>2.5</td>
<td>15.8</td>
<td>2.58</td>
<td>1,894</td>
<td>1,487</td>
<td>7,020</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>38.1</td>
<td>1.5</td>
<td>12.8</td>
<td>2.59</td>
<td>3,001</td>
<td>2,453</td>
<td>11,180</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>37.2</td>
<td>0.9</td>
<td>10.7</td>
<td>2.59</td>
<td>4,707</td>
<td>3,945</td>
<td>17,570</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>36.6</td>
<td>0.6</td>
<td>9.3</td>
<td>2.60</td>
<td>7,333</td>
<td>6,247</td>
<td>27,420</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>36.2</td>
<td>0.4</td>
<td>8.2</td>
<td>2.60</td>
<td>11,376</td>
<td>9,792</td>
<td>42,580</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>35.9</td>
<td>0.2</td>
<td>7.3</td>
<td>2.60</td>
<td>17,599</td>
<td>15,250</td>
<td>65,930</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>35.8</td>
<td>0.2</td>
<td>6.6</td>
<td>2.60</td>
<td>27,179</td>
<td>23,653</td>
<td>101,850</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>35.7</td>
<td>0.1</td>
<td>6.1</td>
<td>2.60</td>
<td>41,924</td>
<td>36,586</td>
<td>157,150</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>35.6</td>
<td>0.1</td>
<td>5.6</td>
<td>2.60</td>
<td>64,620</td>
<td>56,490</td>
<td>242,300</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>35.6</td>
<td>0.0</td>
<td>5.2</td>
<td>2.60</td>
<td>99,550</td>
<td>87,140</td>
<td>373,300</td>
</tr>
<tr>
<td>13</td>
<td>0</td>
<td>35.5</td>
<td>0.0</td>
<td>4.9</td>
<td>2.60</td>
<td>153,320</td>
<td>134,310</td>
<td>575,000</td>
</tr>
<tr>
<td>14</td>
<td>0</td>
<td>35.5</td>
<td>0.0</td>
<td>4.6</td>
<td>2.60</td>
<td>236,090</td>
<td>206,910</td>
<td>885,400</td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>35.5</td>
<td>0.0</td>
<td>4.3</td>
<td>2.60</td>
<td>363,480</td>
<td>318,660</td>
<td>1,363,200</td>
</tr>
</tbody>
</table>

### Table B.1.: Numerical Result 5.2.

**Numerical Result 5.2.: Transition Path With Beveridgian Social Security Under The Assumption Longevity Is Dependent on Human Capital Attainment**

<table>
<thead>
<tr>
<th></th>
<th>τ</th>
<th>Gini(h)</th>
<th>Gini(T)</th>
<th>Gini(U)</th>
<th>T</th>
<th>H</th>
<th>K</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.05</td>
<td>51.6</td>
<td>14.6</td>
<td>49.0</td>
<td>2.42</td>
<td>444</td>
<td>144</td>
<td>1,510</td>
</tr>
<tr>
<td>1</td>
<td>0.05</td>
<td>41.2</td>
<td>6.1</td>
<td>19.1</td>
<td>2.51</td>
<td>767</td>
<td>354</td>
<td>2,700</td>
</tr>
<tr>
<td>2</td>
<td>0.05</td>
<td>33.6</td>
<td>2.6</td>
<td>11.8</td>
<td>2.56</td>
<td>1,310</td>
<td>666</td>
<td>4,650</td>
</tr>
<tr>
<td>3</td>
<td>0.05</td>
<td>28.0</td>
<td>1.2</td>
<td>8.0</td>
<td>2.58</td>
<td>2,183</td>
<td>1,189</td>
<td>7,810</td>
</tr>
<tr>
<td>4</td>
<td>0.05</td>
<td>23.8</td>
<td>0.6</td>
<td>5.7</td>
<td>2.59</td>
<td>3,589</td>
<td>2,037</td>
<td>12,890</td>
</tr>
<tr>
<td>5</td>
<td>0.05</td>
<td>20.5</td>
<td>0.3</td>
<td>4.3</td>
<td>2.60</td>
<td>5,852</td>
<td>3,405</td>
<td>21,070</td>
</tr>
<tr>
<td>6</td>
<td>0.05</td>
<td>17.7</td>
<td>0.2</td>
<td>3.3</td>
<td>2.60</td>
<td>9,493</td>
<td>5,606</td>
<td>34,220</td>
</tr>
<tr>
<td>7</td>
<td>0.05</td>
<td>15.4</td>
<td>0.1</td>
<td>2.6</td>
<td>2.60</td>
<td>15,349</td>
<td>9,149</td>
<td>55,380</td>
</tr>
<tr>
<td>8</td>
<td>0.05</td>
<td>13.4</td>
<td>0.0</td>
<td>2.0</td>
<td>2.60</td>
<td>24,769</td>
<td>14,847</td>
<td>89,430</td>
</tr>
<tr>
<td>9</td>
<td>0.05</td>
<td>11.7</td>
<td>0.0</td>
<td>1.6</td>
<td>2.60</td>
<td>39,923</td>
<td>24,015</td>
<td>144,190</td>
</tr>
<tr>
<td>10</td>
<td>0.05</td>
<td>10.3</td>
<td>0.0</td>
<td>1.3</td>
<td>2.60</td>
<td>64,300</td>
<td>38,762</td>
<td>232,280</td>
</tr>
<tr>
<td>11</td>
<td>0.05</td>
<td>9.0</td>
<td>0.0</td>
<td>1.1</td>
<td>2.60</td>
<td>103,500</td>
<td>62,480</td>
<td>374,000</td>
</tr>
<tr>
<td>12</td>
<td>0.05</td>
<td>7.9</td>
<td>0.0</td>
<td>0.9</td>
<td>2.60</td>
<td>166,600</td>
<td>100,640</td>
<td>601,900</td>
</tr>
<tr>
<td>13</td>
<td>0.05</td>
<td>6.9</td>
<td>0.0</td>
<td>0.7</td>
<td>2.60</td>
<td>268,100</td>
<td>162,030</td>
<td>968,600</td>
</tr>
<tr>
<td>14</td>
<td>0.05</td>
<td>6.0</td>
<td>0.0</td>
<td>0.6</td>
<td>2.60</td>
<td>431,300</td>
<td>260,770</td>
<td>1,558,500</td>
</tr>
<tr>
<td>15</td>
<td>0.05</td>
<td>5.3</td>
<td>0.0</td>
<td>0.5</td>
<td>2.60</td>
<td>693,800</td>
<td>419,610</td>
<td>2,507,300</td>
</tr>
</tbody>
</table>
### Table A.B. con.: Numerical Result 5.2.

#### Numerical Result 5.2.i.: Transition Path With Beveridgian Social Security Under The Assumption

**Longevity is Dependent on Human Capital Attainment**

<table>
<thead>
<tr>
<th>$t$</th>
<th>$\tau$</th>
<th>Gini(h)</th>
<th>Gini(T)</th>
<th>Gini(U)</th>
<th>T</th>
<th>H</th>
<th>K</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.1</td>
<td>51.6</td>
<td>14.6</td>
<td>31.2</td>
<td>2.42</td>
<td>444</td>
<td>144</td>
<td>1,510</td>
</tr>
<tr>
<td>1</td>
<td>0.1</td>
<td>38.8</td>
<td>5.4</td>
<td>15.8</td>
<td>2.51</td>
<td>781</td>
<td>272</td>
<td>2,670</td>
</tr>
<tr>
<td>2</td>
<td>0.1</td>
<td>29.5</td>
<td>2.1</td>
<td>9.2</td>
<td>2.56</td>
<td>1,319</td>
<td>511</td>
<td>4,560</td>
</tr>
<tr>
<td>3</td>
<td>0.1</td>
<td>22.9</td>
<td>0.9</td>
<td>5.9</td>
<td>2.58</td>
<td>2,181</td>
<td>906</td>
<td>7,590</td>
</tr>
<tr>
<td>4</td>
<td>0.1</td>
<td>18.1</td>
<td>0.4</td>
<td>3.9</td>
<td>2.59</td>
<td>3,565</td>
<td>1,544</td>
<td>12,460</td>
</tr>
<tr>
<td>5</td>
<td>0.1</td>
<td>14.4</td>
<td>0.2</td>
<td>2.8</td>
<td>2.60</td>
<td>5,783</td>
<td>2,567</td>
<td>20,260</td>
</tr>
<tr>
<td>6</td>
<td>0.1</td>
<td>11.6</td>
<td>0.1</td>
<td>2.0</td>
<td>2.60</td>
<td>9,339</td>
<td>4,208</td>
<td>32,770</td>
</tr>
<tr>
<td>7</td>
<td>0.1</td>
<td>9.3</td>
<td>0.1</td>
<td>1.4</td>
<td>2.60</td>
<td>15,038</td>
<td>6,839</td>
<td>52,820</td>
</tr>
<tr>
<td>8</td>
<td>0.1</td>
<td>7.5</td>
<td>0.0</td>
<td>1.1</td>
<td>2.60</td>
<td>24,175</td>
<td>11,057</td>
<td>84,950</td>
</tr>
<tr>
<td>9</td>
<td>0.1</td>
<td>6.1</td>
<td>0.0</td>
<td>0.8</td>
<td>2.60</td>
<td>38,819</td>
<td>17,818</td>
<td>136,460</td>
</tr>
<tr>
<td>10</td>
<td>0.1</td>
<td>4.9</td>
<td>0.0</td>
<td>0.6</td>
<td>2.60</td>
<td>62,291</td>
<td>28,655</td>
<td>219,020</td>
</tr>
<tr>
<td>11</td>
<td>0.1</td>
<td>4.0</td>
<td>0.0</td>
<td>0.5</td>
<td>2.60</td>
<td>99,900</td>
<td>46,030</td>
<td>351,400</td>
</tr>
<tr>
<td>12</td>
<td>0.1</td>
<td>3.2</td>
<td>0.0</td>
<td>0.3</td>
<td>2.60</td>
<td>160,200</td>
<td>73,870</td>
<td>563,500</td>
</tr>
<tr>
<td>13</td>
<td>0.1</td>
<td>2.6</td>
<td>0.0</td>
<td>0.3</td>
<td>2.60</td>
<td>256,900</td>
<td>118,500</td>
<td>903,500</td>
</tr>
<tr>
<td>14</td>
<td>0.1</td>
<td>2.1</td>
<td>0.0</td>
<td>0.2</td>
<td>2.60</td>
<td>411,800</td>
<td>190,030</td>
<td>1,448,500</td>
</tr>
<tr>
<td>15</td>
<td>0.1</td>
<td>1.7</td>
<td>0.0</td>
<td>0.2</td>
<td>2.60</td>
<td>660,200</td>
<td>304,690</td>
<td>2,322,000</td>
</tr>
</tbody>
</table>

#### Numerical Result 5.2.ii.: Transition Path With Beveridgian Social Security Under The Assumption

**Longevity Is Dependent on Human Capital Attainment**

<table>
<thead>
<tr>
<th>$t$</th>
<th>$\tau$</th>
<th>Gini(h)</th>
<th>Gini(T)</th>
<th>Gini(U)</th>
<th>T</th>
<th>H</th>
<th>K</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.15</td>
<td>51.6</td>
<td>14.6</td>
<td>29.2</td>
<td>2.42</td>
<td>444</td>
<td>144</td>
<td>1,510</td>
</tr>
<tr>
<td>1</td>
<td>0.15</td>
<td>37.4</td>
<td>5.1</td>
<td>14.3</td>
<td>2.51</td>
<td>773</td>
<td>216</td>
<td>2,580</td>
</tr>
<tr>
<td>2</td>
<td>0.15</td>
<td>27.2</td>
<td>1.9</td>
<td>8.1</td>
<td>2.56</td>
<td>1,264</td>
<td>395</td>
<td>4,280</td>
</tr>
<tr>
<td>3</td>
<td>0.15</td>
<td>20.1</td>
<td>0.8</td>
<td>5.0</td>
<td>2.58</td>
<td>2,032</td>
<td>680</td>
<td>6,920</td>
</tr>
<tr>
<td>4</td>
<td>0.15</td>
<td>15.1</td>
<td>0.4</td>
<td>3.2</td>
<td>2.59</td>
<td>3,230</td>
<td>1,126</td>
<td>11,050</td>
</tr>
<tr>
<td>5</td>
<td>0.15</td>
<td>11.4</td>
<td>0.2</td>
<td>2.2</td>
<td>2.60</td>
<td>5,099</td>
<td>1,823</td>
<td>17,480</td>
</tr>
<tr>
<td>6</td>
<td>0.15</td>
<td>8.7</td>
<td>0.1</td>
<td>1.5</td>
<td>2.60</td>
<td>8,015</td>
<td>2,911</td>
<td>27,520</td>
</tr>
<tr>
<td>7</td>
<td>0.15</td>
<td>6.6</td>
<td>0.0</td>
<td>1.0</td>
<td>2.60</td>
<td>12,562</td>
<td>4,608</td>
<td>43,180</td>
</tr>
<tr>
<td>8</td>
<td>0.15</td>
<td>5.1</td>
<td>0.0</td>
<td>0.7</td>
<td>2.60</td>
<td>19,654</td>
<td>7,255</td>
<td>67,600</td>
</tr>
<tr>
<td>9</td>
<td>0.15</td>
<td>3.9</td>
<td>0.0</td>
<td>0.5</td>
<td>2.60</td>
<td>30,714</td>
<td>11,383</td>
<td>105,680</td>
</tr>
<tr>
<td>10</td>
<td>0.15</td>
<td>3.0</td>
<td>0.0</td>
<td>0.4</td>
<td>2.60</td>
<td>47,962</td>
<td>17,822</td>
<td>165,100</td>
</tr>
<tr>
<td>11</td>
<td>0.15</td>
<td>2.3</td>
<td>0.0</td>
<td>0.3</td>
<td>2.60</td>
<td>746,600</td>
<td>27,860</td>
<td>257,700</td>
</tr>
<tr>
<td>12</td>
<td>0.15</td>
<td>1.8</td>
<td>0.0</td>
<td>0.2</td>
<td>2.60</td>
<td>1,168,200</td>
<td>43,520</td>
<td>402,200</td>
</tr>
<tr>
<td>13</td>
<td>0.15</td>
<td>1.4</td>
<td>0.0</td>
<td>0.1</td>
<td>2.60</td>
<td>1,822,400</td>
<td>67,950</td>
<td>627,500</td>
</tr>
<tr>
<td>14</td>
<td>0.15</td>
<td>1.1</td>
<td>0.0</td>
<td>0.1</td>
<td>2.60</td>
<td>2,842,800</td>
<td>106,030</td>
<td>978,800</td>
</tr>
<tr>
<td>15</td>
<td>0.15</td>
<td>0.8</td>
<td>0.0</td>
<td>0.1</td>
<td>2.60</td>
<td>4,434,200</td>
<td>165,440</td>
<td>1,526,800</td>
</tr>
<tr>
<td>$t$</td>
<td>$\tau$</td>
<td>$\text{Gini}(h)$</td>
<td>$\text{Gini}(T)$</td>
<td>$\text{Gini}(U)$</td>
<td>$T$</td>
<td>$H$</td>
<td>$K$</td>
<td>$Y$</td>
</tr>
<tr>
<td>-----</td>
<td>--------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>0</td>
<td>0.16</td>
<td>51.6</td>
<td>14.6</td>
<td>29.0</td>
<td>2.42</td>
<td>444</td>
<td>144</td>
<td>1,510</td>
</tr>
<tr>
<td>1</td>
<td>0.16</td>
<td>37.2</td>
<td>5.1</td>
<td>14.1</td>
<td>2.51</td>
<td>769</td>
<td>207</td>
<td>2,560</td>
</tr>
<tr>
<td>2</td>
<td>0.16</td>
<td>26.8</td>
<td>1.9</td>
<td>7.9</td>
<td>2.56</td>
<td>1,249</td>
<td>375</td>
<td>4,210</td>
</tr>
<tr>
<td>3</td>
<td>0.16</td>
<td>19.7</td>
<td>0.8</td>
<td>4.8</td>
<td>2.58</td>
<td>1,992</td>
<td>641</td>
<td>6,760</td>
</tr>
<tr>
<td>4</td>
<td>0.16</td>
<td>14.6</td>
<td>0.4</td>
<td>3.1</td>
<td>2.59</td>
<td>3,144</td>
<td>1,054</td>
<td>10,710</td>
</tr>
<tr>
<td>5</td>
<td>0.16</td>
<td>11.0</td>
<td>0.2</td>
<td>2.1</td>
<td>2.60</td>
<td>4,927</td>
<td>1,694</td>
<td>16,830</td>
</tr>
<tr>
<td>6</td>
<td>0.16</td>
<td>8.3</td>
<td>0.1</td>
<td>1.4</td>
<td>2.60</td>
<td>7,688</td>
<td>2,687</td>
<td>26,300</td>
</tr>
<tr>
<td>7</td>
<td>0.16</td>
<td>6.3</td>
<td>0.0</td>
<td>1.0</td>
<td>2.60</td>
<td>11,963</td>
<td>4,223</td>
<td>40,960</td>
</tr>
<tr>
<td>8</td>
<td>0.16</td>
<td>4.8</td>
<td>0.0</td>
<td>0.7</td>
<td>2.60</td>
<td>18,580</td>
<td>6,002</td>
<td>63,660</td>
</tr>
<tr>
<td>9</td>
<td>0.16</td>
<td>3.6</td>
<td>0.0</td>
<td>0.5</td>
<td>2.60</td>
<td>28,824</td>
<td>10,285</td>
<td>98,810</td>
</tr>
<tr>
<td>10</td>
<td>0.16</td>
<td>2.8</td>
<td>0.0</td>
<td>0.3</td>
<td>2.60</td>
<td>44,683</td>
<td>15,986</td>
<td>153,210</td>
</tr>
<tr>
<td>11</td>
<td>0.16</td>
<td>2.1</td>
<td>0.0</td>
<td>0.2</td>
<td>2.60</td>
<td>69,230</td>
<td>24,810</td>
<td>237,400</td>
</tr>
<tr>
<td>12</td>
<td>0.16</td>
<td>1.6</td>
<td>0.0</td>
<td>0.2</td>
<td>2.60</td>
<td>107,240</td>
<td>384,700</td>
<td>367,800</td>
</tr>
<tr>
<td>13</td>
<td>0.16</td>
<td>1.2</td>
<td>0.0</td>
<td>0.1</td>
<td>2.60</td>
<td>166,070</td>
<td>596,300</td>
<td>569,600</td>
</tr>
<tr>
<td>14</td>
<td>0.16</td>
<td>0.9</td>
<td>0.0</td>
<td>0.1</td>
<td>2.60</td>
<td>257,150</td>
<td>923,700</td>
<td>882,100</td>
</tr>
<tr>
<td>15</td>
<td>0.16</td>
<td>0.7</td>
<td>0.0</td>
<td>0.1</td>
<td>2.60</td>
<td>398,140</td>
<td>1,430,600</td>
<td>1,365,700</td>
</tr>
</tbody>
</table>

**Numerical Result 5.2.iii.: Transition Path With Beverdigan Social Security Under The Assumption Longevity Is Dependent on Human Capital Attainment**

<table>
<thead>
<tr>
<th>$t$</th>
<th>$\tau$</th>
<th>$\text{Gini}(h)$</th>
<th>$\text{Gini}(T)$</th>
<th>$\text{Gini}(U)$</th>
<th>$T$</th>
<th>$H$</th>
<th>$K$</th>
<th>$Y$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.17</td>
<td>51.6</td>
<td>14.6</td>
<td>29.0</td>
<td>2.42</td>
<td>444</td>
<td>144</td>
<td>1,510</td>
</tr>
<tr>
<td>1</td>
<td>0.17</td>
<td>37.1</td>
<td>5.0</td>
<td>14.1</td>
<td>2.51</td>
<td>767</td>
<td>203</td>
<td>2,550</td>
</tr>
<tr>
<td>2</td>
<td>0.17</td>
<td>26.6</td>
<td>1.9</td>
<td>7.9</td>
<td>2.56</td>
<td>1,240</td>
<td>366</td>
<td>4,170</td>
</tr>
<tr>
<td>3</td>
<td>0.17</td>
<td>19.5</td>
<td>0.8</td>
<td>4.8</td>
<td>2.58</td>
<td>1,971</td>
<td>622</td>
<td>6,670</td>
</tr>
<tr>
<td>4</td>
<td>0.17</td>
<td>14.4</td>
<td>0.4</td>
<td>3.1</td>
<td>2.59</td>
<td>3,099</td>
<td>1,019</td>
<td>10,540</td>
</tr>
<tr>
<td>5</td>
<td>0.17</td>
<td>10.8</td>
<td>0.2</td>
<td>2.0</td>
<td>2.60</td>
<td>4,838</td>
<td>1,633</td>
<td>16,490</td>
</tr>
<tr>
<td>6</td>
<td>0.17</td>
<td>8.1</td>
<td>0.1</td>
<td>1.4</td>
<td>2.60</td>
<td>7,521</td>
<td>2,579</td>
<td>25,680</td>
</tr>
<tr>
<td>7</td>
<td>0.17</td>
<td>6.1</td>
<td>0.0</td>
<td>1.0</td>
<td>2.60</td>
<td>11,658</td>
<td>4,040</td>
<td>39,850</td>
</tr>
<tr>
<td>8</td>
<td>0.17</td>
<td>4.6</td>
<td>0.0</td>
<td>0.7</td>
<td>2.60</td>
<td>18,038</td>
<td>6,291</td>
<td>61,690</td>
</tr>
<tr>
<td>9</td>
<td>0.17</td>
<td>3.5</td>
<td>0.0</td>
<td>0.5</td>
<td>2.60</td>
<td>27,876</td>
<td>9,764</td>
<td>95,380</td>
</tr>
<tr>
<td>10</td>
<td>0.17</td>
<td>2.7</td>
<td>0.0</td>
<td>0.3</td>
<td>2.60</td>
<td>43,046</td>
<td>15,119</td>
<td>147,330</td>
</tr>
<tr>
<td>11</td>
<td>0.17</td>
<td>2.0</td>
<td>0.0</td>
<td>0.2</td>
<td>2.60</td>
<td>66,440</td>
<td>23,380</td>
<td>227,400</td>
</tr>
<tr>
<td>12</td>
<td>0.17</td>
<td>1.5</td>
<td>0.0</td>
<td>0.2</td>
<td>2.60</td>
<td>1,025,100</td>
<td>36,110</td>
<td>350,900</td>
</tr>
<tr>
<td>13</td>
<td>0.17</td>
<td>1.2</td>
<td>0.0</td>
<td>0.1</td>
<td>2.60</td>
<td>1,581,300</td>
<td>55,740</td>
<td>541,400</td>
</tr>
<tr>
<td>14</td>
<td>0.17</td>
<td>0.9</td>
<td>0.0</td>
<td>0.1</td>
<td>2.60</td>
<td>2,439,100</td>
<td>86,020</td>
<td>835,100</td>
</tr>
<tr>
<td>15</td>
<td>0.17</td>
<td>0.7</td>
<td>0.0</td>
<td>0.1</td>
<td>2.60</td>
<td>3,761,700</td>
<td>132,710</td>
<td>1,288,000</td>
</tr>
</tbody>
</table>
GREGORY GILPIN

Department of Economics – Indiana University  Email: gagilpin@indiana.edu
Wylie Hall 105  Phone: 812-855-9531 (work)
100 S. Woodlawn  812 339-5883 (home)
Bloomington, IN 47405-7104  Fax: 812-855-3736
Web: http://mypage.iu.edu/~gagilpin

Education

Ph.D., Economics, Indiana University, 2009
M.A., Economics, University of Kansas, 2004
B.S., Mathematics, Economics minor, Brigham Young University - Hawaii, 2002

Dissertation

Committee: Michael Kaganovich (Chair), William E. Becker, Gerhard Glomm, Rusty Tchernis

Fields of Specialization

Primary: Macroeconomics, Public Economics, Applied Econometrics
Secondary: Growth Theory, Monetary Economics, Computational Methods

Teaching Interests (undergraduate and graduate)

Macroeconomics, Microeconomics, Econometrics, Public Finance, Labor Economics, Statistics

Research Interests

Macroeconomic Aspects of Public Policy, Public Economics, Applied Econometrics
(JEL classification: C11; E60; H52; I22; O4)

Working Papers

• Self-selection, Unions, and Teacher Attrition
• Unequal Longevity and Social Security: The Implications for Human Capital, Growth and Distribution
• The Quantity and Quality of Teachers: A Dynamic Trade-off (with Michael Kaganovich)
• The Effect of Wage Frictions on Teacher Quality: Are Math and Science Teachers Affected Differently? (with Chris Ferguson; winner of the Bill and Melinda Gates Foundation Teacher Quality Paper Competition)

Presentations at Conferences and Seminars

Midwest Macro Meeting, University of Pennsylvania, 2008
Midwest Econometrics Group, University of Kansas, 2008
Missouri Economic Conference, University of Missouri, Columbia, 2008, awarded travel grant
Jordan River Conference, Indiana University, 2008
Workshop in Public and Environmental Affairs, Indiana University, 2008
Workshop in Discipline-based Scholarship in Education, Indiana University, 2008
Jordan River Conference, Indiana University, 2007

Research Experience

Consultant, Electrical & Utilities Safety Association, Mississauga, Ontario, 2005 to present
Research Intern, Economic Research Department, FRB of Kansas City, 2004
Research Assistant, Policy Research Institute, University of Kansas, 2003-04
Research Assistant, School of Business, Brigham Young University, 2000-02

Teaching Experience

Associate Instructor (full teaching responsibilities)

Public Finance (undergraduate) Spring 2008, Spring 2009 Indiana University
Statistical Analysis for Bus. & Econ. Spring 2006 (3 lab sections) Indiana University
Principles of Microeconomics Fall 2008, Summer 2008, Indiana University
Fall 2007, Spring 2007 Indiana University
Fall 2006, Fall 2005 Indiana University
Summer 2005 IUPUI

Discussion Leader (conduct review sessions)

Managerial Economics Spring 2004 (3 sections) Univ. of Kansas
Principles of Macroeconomics Spring 2003 (2 sections) Univ. of Kansas
Principles of Microeconomics Fall 2002 (3 sections) Univ. of Kansas

Teaching Assistant (grading, advising, and periodic lecturing)

Econometric Theory and Practice II Spring 2005 Indiana University
Econometric Theory and Practice I Fall 2004 Indiana University
Intermediate Microeconomic Theory Fall 2004 Indiana University
Principles of Statistics I (Math Dept.) Spring 2002 BYU - Hawaii
Calculus I (Math Dept.) Fall 2001, Fall 2000 BYU - Hawaii
Calculus II (Math Dept.) Spring 2001 BYU - Hawaii

Honors & Awards

• Help or Hindrance: The Impact of Teacher Rules, Roles, and Rights on Teacher Quality, Research Competition, Bill and Melinda Gates Foundation, $5,000, 2008
• W. Phillip Saunders Teaching Award, IU Economics, 2008
• Daniel J. Duesterberg Award for Research in Public Economics, IU Economics, 2008
• Discipline Based Scholarship in Education Associate (Fellowship), Spencer Foundation, $2,000, 2007
• College of Arts and Sciences Graduate Fellowship, IU, $5,000, 2004
• Employee Excellence Research Award, Federal Reserve Bank of Kansas City, 2004
• Math & Science Division Academic Scholarship, Brigham Young University - Hawaii, $2,000, 2002
• Carl C. Conway Academic Scholarship, Carl C. Conway Foundation, $36,500, 1996-97, 1999-2002

Reports and Manuals


Associations  Member of Omicron Delta Epsilon, Southern Economic Association, Canadian Economic Association

Languages  English, Spanish

Software  Matlab, SAS, STATA

Nationalities  American, Canadian
References

Teaching References

William E. Becker
Professor of Economics
Indiana University
100 S. Woodlawn
Bloomington, IN 47405
Phone: (812) 855-3577
E-mail: beckerw@indiana.edu

James K. Self
Lecturer of Economics
Indiana University
100 S. Woodlawn
Bloomington, IN 47405
Phone: (812) 855-7590
E-mail: jkselI@indiana.edu

Research References

Michael Kaganovich
Professor of Economics and
Graduate Director
Indiana University
100 S. Woodlawn
Bloomington, IN 47405
Phone (812) 855-6967
E-mail: mkaganov@indiana.edu

William E. Becker
Professor of Economics
Indiana University
100 S. Woodlawn
Bloomington, IN 47405
Phone: (812) 855-3577
E-mail: beckerw@indiana.edu

Gerhard Glomm
Professor of Economics and
Department Chair
Indiana University
100 S Woodlawn
Bloomington, IN 47405
Phone: (812) 855-7256
E-mail: gglomm@indiana.edu

Rusty Tchernis
Assistant Professor of Economics
Indiana University
100 S Woodlawn
Bloomington, IN 47405
Phone: (812) 855-7791
E-mail: rtcherni@indiana.edu