Generational Accounts for Public Education in the US: Historical and Projected

Wenlan Qian

June 5, 2002

Abstract

This project aims to derive the net present value (NPV) of public educational benefits of birth cohorts from 1850 to 2000 in the United States. It is calculated in three steps. First, we obtain the national average public expenditures for education per person age x in year t from 1850 to present and make projections about the expenditure till 2090 based on a fixed 1.8% growth rate after year 1990. Next, we derived the age shape of taxation, i.e. the national average tax payments for public education by persons age x in year t from 1850 to 2090. Finally, we converted the period measure of public educational benefits and tax payments into those for each cohort member who was born between 1850 and 2000. The difference between the lifetime public educational benefits and tax payments for each cohort member is the NPV associated with the particular cohort. We also discounted the education expenditures and tax payments with both the conventional 3% and the historical interest rate series, as well as weighted by survival rates. In addition, we derived the lifetime earnings for all those cohorts and therefore derive a time series of the NPV as a percentage in the lifetime earnings. Results show that the cohorts born around 1938 incur the biggest losses: such a cohort member on average has a NPV of over -$11,000 (1999 constant dollars) when discounted at 3%. In addition, their loss in the net education benefits amounts to 4% in their lifetime earnings.

Part I
Main Assumptions

However, to realize these goals, we encounter many practical problems: The population, education expenditure and taxes data mainly come from the census from IPUMS (**I**ntegrated **P**ublic **U**se **M**icrodata **S**eries). We only have 13 census years between the periods 1850 to 1990: the 1890 and 1930 census data are unavailable from IPUMS. In addition, data on education expenditure and tax payments beyond year 1990 are unavailable. In order to accomodate the lack of data, we make a few assumptions about the education expenditure and tax payments. The biggest assumption in this project is that, due to the incontinuity of the census data, we use interpolation and smoothing to obtain the single year data. Furthermore, for derivation of education expenditure profiles, tax profiles and population counts, we each have several assumptions:

Assumptions for education expenditure calculation

- In the census data we get from IPUMS, there is no direct observation of the educational expenditures.To calculate the public expenditure per capita, we use a time series of the public expenditure per pupil (either directly available or derived by the total expenditures and total enrolled students, from *the Historical Statistics of the United States* and *Digest of Education Statistics*) to time the public enrollment rate, which is in turn derived by multiplying enrollment rate by proportion of public schools. And the same is for public versus private enrollment.

- In some censuses, day care schools and nursery school enrollments are also reported. To keep consistency, we eliminate all enrollments under age 5.

- Ideally, we want three categories for the educational expenditure side: elementary, high school and college. Given the expenditure data that we have, we could only have two, with elementary and high school combined.

- There is no source that specifies grade advancement by age. Therefore, we make assumptions to classify school-age populations into the two categories (elementary and secondary versus college). One more thing to pay attention to is that kindergarten enrollment and expenditures are included in the elementary data for some or all the years. Since they account for a trivial part of the expenditure profile, their effect is negligible.

- To make projections about education expenditure profile after year 1990, we assume that the public education expenditure will grow at the same speed as the projected labor productivity growth rate, 1.8%.[1]

Assumptions for age shape of taxation

- Public education is funded by property tax. There could be other sources that funds public education (e.g. income tax), but given the data constraint, we base our calculation solely on property taxes.

- The tax that pays for public education is proprotional to the current value of the property.

- In the census, this value is reported by respondents who own their own home. Renters report their average monthly rent which we assume is proprotional to the value of the property. We use census data from 1940 to 1990 to derive the age profile of home value for heads who own their homes and the age profile of monthly rent for heads who pay rent.

[1] The 1.8% is quoted from the paper done by Ronald Lee and Ryan Edwards, *The Fiscal Impact of Population Aging in the US: Assessing the Uncertainties*

- Data from the BEA (**B**ureau of the **E**conomic **A**nalysis) yields an accurate estimate of the aggregate value of residential housing by tenure: owned and rented from 1925 to 1990. We use this data to adjust the levels of the two age profiles derived by the census.
- About 70% of taxes on rental properties are paid by renters in teh form of higher rents.
- The age profile of home owned by landlords is the same shape as the age profile of owned-home values.
- For a given period/year, the sum of total taxes paid (for education) equal the sum of total public education expenditure. This is the final step to adjust the level of the tax profile. After we obtain the final age shape of taxation, we adjusted their levels with the aggregate education expenditure in each period.
- We assume people under age 15 do not pay property taxes and make their tax payments 0.

Other Assumptions

- The 1850 and 1860 census data we get from IPUMS do not include slaves. To make the population count consistent for all years, we have to adjust the 1850 and 1860 population data. Historical Statistics have a table that has the counts of slave population by age (five or ten years' age group). We read in the numbers, evenly distribute them to every single age and add to IPUMS population counts. Since these are slave population, we assume that none of them attend school or own property.
- In the last step of calculating NPV, we need to discount the national public education expenditure/tax payments per person with historical interest rates. For the missing values ($<$ 1890 and $>$ 1997), we have to use the average of the closest ten years' interest rates and apply it through out those years.

Part II

Data Collection and Age Profiles from 13 Censuses

We use IPUMS and get 13 micro-level census data from 1850 to 1990. After data extraction and manipulation, we get 13 weighted age profiles, from which we have population counts by age, enrollment rates by age, occupation status by age (teacher vs. non-teacher), income by age (for 1850 through 1870 we have real property by age instead), proportion that owns/rents houses by age and proportion of household heads by age.

General assumptions and technical procedures: The data are a 1% sample of the census; extract file type is rectangular; command file format is SAS. Select variables are as follows:

- Basic demographic variables: age, sex
- Education variables:
  - All 13 censuses data extracts contain variable school, which describes enrollment.
  - For public and private school classification, we have the variable schltype since 1960 (available in total in four censuses); Since 1940, we further add an education-related variable: educrec, i.e. the educational attainment.
- Income variables: In addition, we also need income for calculating tax contribution per person by age. The best information in the census data is the income-related variables. However, they vary for each census data. We just download the different income-related data for now and will adjust it later.
  - 1850: realprop (real estate value)
  - 1860-1870: realprop, persprop (value of personal estate)
  - 1880: no income variables available from IPUMS
  - 1900 - 1920 (except 1950): ownershp (ownership of dwelling); relate (relationship to household head); mortgage (mortgage status);
  - 1940: ownershp; relate; value (house or property value); rent (monthly contract rent); fwage1 (family wage and salary income)
  - 1950: ftotinc (total family income); fwage1; fbusiness (business income of other family members); ftothine (other income of other family members)
  - 1960-1970: ownershp; relate; value; rent; fotoinc
  - 1980-1990: ownershp; relate; value; rent; ftotinc;
    hhincome (total household income)
- weight variable: Since all census extracts are 1% sample size of the entire population, we also get the weight variable to adjust the result. For all 13 censuses, we retrieve the weight variable perwt (personal weight) to adjust for most of the variables. One exception is the year 1950, where we need one more weight variable, slwt (sample line weight), to adjust for its income variables.

Data Manipulation We keep the 13 compressed census data extracts in thirteen directories after downloading them. In order to getage profiles of each census data, we use SAS programming language.

For some continuous variables like age, educrec, perwt, slwt, realprop, persprop, value, rent, ftotinc, fwage1, fothine and hhincome, we retain the value of the variables as they are. For the school variable, we turn it into a category variable by classifying those currently in school (whatever grade or status) as 1 and the remaining as 0. Thus we can get the enrollment rate by averaging over the whole age group weighted by the weight variable (usually perwt).

The same strategy is used on variables such as ownershp, relate, occ, ind, schltype, through which we can get the proportion of the age group owning houses, proportion that are heads of the households, proportion in public school and percentage of age groups that are teachers.

We then sorted the data by age, summarize the data by age (by calculating the means of the variable value over that single age population) and put the weighted results into ASCII files weighted.age.profiles.*datayear*. Note that we also put the count of the population at every age in the age profiles so that at the end we have the population by every single age for each of the thirteen census years.

Other immediate results of the SAS work are thirteen enrollment rates by single year of age, thirteen proportion of the population that are teachers by single year of age, eight proportion of the population owning homes by single year of age (starting year 1900), eight proportion of the population who are household heads by single year of age, four proportion of population in public schools by single year of age (starting 1960).

Note that all we have done so far are mostly dealing with educational data; the tax data manipulation cannot be completed with the available data from IPUMS.

Part III
NPV Calculations

The NPV is a summary measure of the stream of benefits less taxes received by a member of a birth cohort. It is survival weighted, discounted sum of the costs of education minus the taxes for education for each birth cohort. In order to calculate the NPV of education for birth cohorts, we need four elements:

A matrix of per capita educational costs by age and year;
A matrix of per capita educational taxes by age and year;
A survival rate matrix for each birth cohort;
A time series of the discount rates

1 Matrix of Educational Expenditures by Age and Year

First of all, we observe that the enrollment rates above age 35 are low and then only count educational costs till age 35. We begin by constructing a matrix of per capita educational costs by age and year. That is, we want to know the average expenditure per person aged x in year t. To calculate NPV, we need to know:

1.1 public expenditures per student enrolled

Ideally, we would want three vectors: public expenditures per college student; public expenditures per high school student; public expenditures per elementary student. However, due to data availability, we can only have two vectors; elementary and secondary student expenditures are combined into one vector.

To calculate the public expenditures per student for every year, we divide the total public expenditures by public school students, both of which are available from the Digest of Education Statistics(on a ten-year basis). We interpolate them to get per student (either elementary and secondary or higher education) for each year. However, the earliest elementary and secondary per pupil expenditure we have is 1870 and we have 1930 as our earliest year for college per student expenditure. We use the available GNP per capita to adjust the level of college per capita expenditure to obtain the values until year 1870 and then assume that the 1870's values apply to the previous years.

One more thing to pay attention to is the kindergarten issue since its enrollment and expenditures are already included in elementary data for some or all the years. However, it should account for a very small portion of the expenditure and thus its effect is neglible.

The last step in this part is to convert all the expenditures into constant dollars. We would want to have a time series of GDP deflators. We get the values between 1940 and 1999 from the website *Budget of the U.S. Government.*[2] For previous years, we use the implicit price index 1869-1970 from the *Historical Statistics of the United States.* We convert both series by making year 1970's value as the reference and plot them on the same graph. The two lines match quite well. So we just combine the two series (values from 1869 to 1939 from implicit price index and data beyond 1940 from GDP deflator) and get a single series of deflators between 1869 and 1999. We use the same series to make adjustments throughout the project and all the expenditure, taxes as well as NPV are represented in constant 1999 dollars.

[2] Section 10–Gross Domestic Product and Implicit Outlay Deflators, Table 10.1–Gross Domestic Product and Deflators Used in the Historical Tables: 1940-2005,http://w3.access.gpo.gov/usbudget/fy2001/hist.html#h10

1.2 Matrix of public school enrollment by age a and year t

From the census data we get previously (IPUMS), we have total enrollment rates by single year of age for each of the 13 censuses. In order to get public school enrollment rates, we need proportion of public schools by year. As to distinguish between public and private schools for every year, we use the interpolated public and private school enrollment numbers for each year that we get from the *Digest* to infer the percent of all enrolled students in elementary, secondary schools or colleges that are in public schools.

In addition, under the conventional assumption that all students aged under 13 are in elementary schools, those between 14 and 17 are in high school, and all 18 to 35 year-old students are in college, the total public college students for each of the 13 censuses calculated from IPUMS do not match the numbers in the *Educational Digest of Statistics.* For most census years, the assumption that students go to college at exactly age 18 tends to overestimate the total college students enrolled.

For correction, we use the enrollment data from *the Digest of Educational Statistics*, which distinguishes the elementary, secondary, and college enrollment from 1870 to 1990. For each census year starting 1870, we try to determine the age of going to college by the method of matching. However, we are still assuming that school advancement occurs at exactly one age point: no overlapping is considered. The new advancement age from elem/second schools to colleges are as follows:

1870	1880	1900	1910	1920	1940	1950	1960	1970	1980	1990
21	21	20	20	19	19	20	19	18	18	19

Thus, multiplying the enrollment by age by proportion of public schools, adjusted by the grade advancement, we get a matrix of public school enrollment by age and year.

With these data, we can calculate the matrix of public school educational costs per capita by age x and time t.

$$\mathbf{C}(a,t) = e(a,t)s(t)c(t), \forall t \in [1850, 1990] \tag{1}$$

For each year t and age a, $\mathbf{C}$(a,t) denotes the matrix of public school educational costs per capita, e(a,t) is the total enrollment rate, s is the proportion of public schools and c(t) is the average expenditure per student. We have to consider part-time enrollment, which is particularly true among people at older ages. Thus we ideally need the time-series of part-time enrollment for all ages. Unfortunately, we don't have that data. We graphed the available part-time enrollment (mainly after year 1960) and made comparisons, from which we found that the part-time enrollment age profile does not change much over time. Therefore, we use a constant age profile(in 1994) and assume that those part-time students are 50% enrolled(i.e. discount the enrollment rate by multiplying the proportion that is part-time enrolled by 1/2). The final public education

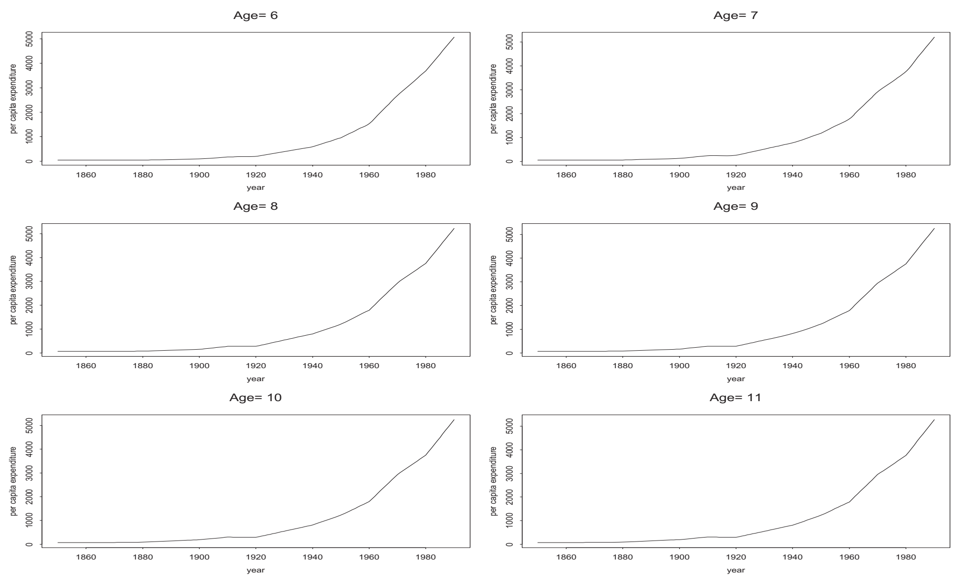

Figure 1: Per Capiture Public Education Expenditure by Time and Age(6-11)

expenditure per capita by age and time is shown in Figure 1(We include only ages from 6 to 23 in the figure, which are the ages we are most concerned with).

Furthermore, for cohort NPV between 1850 and 2000, we will need period educational costs till year 2090 (we trace the population till age 90 (and over)). To be consistent with other ongoing work, we use the predicted labor productivity growth rate 1.8% as the proxy for the educational expenditure growth rate. With that fixed rate, we predict the public education expenditure after year 1990 based on the value in that year. So:

$$\mathbf{C}(a,t) = e(a,t)s(t)c(t), \forall t \in [1850, 1990] \quad (2)$$

$$= e(a,1990)s(1990)exp(0.018t), \forall t \in (1990, 2090] \quad (3)$$

Finally, we have a complete series of education expenditure profile from year 1850 to 2090.

2 Matrix of educational taxes by age and year

2.1 Assmuptions:

1. The basic assumption is that total education taxes equal these total expenditures.

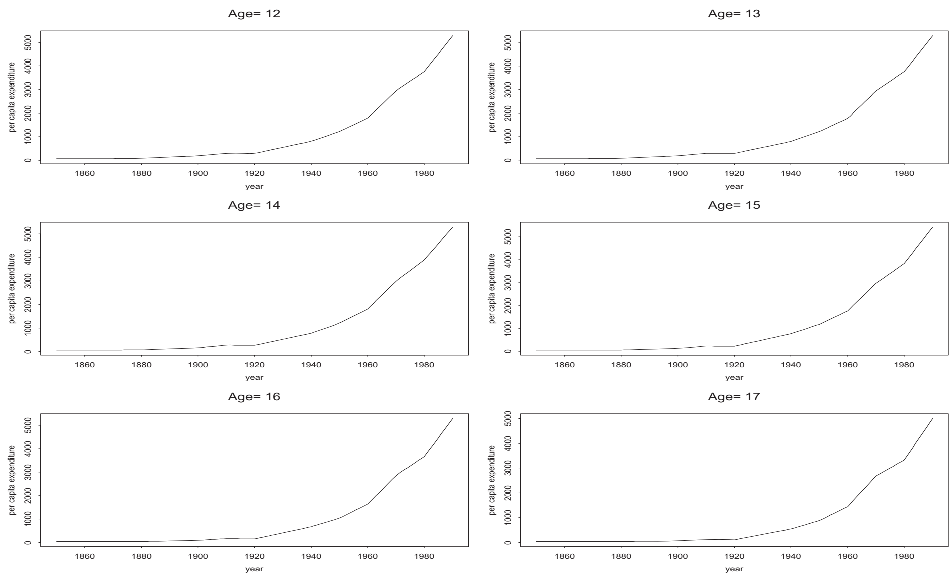

Figure 2: Per Capiture Public Education Expenditure by Time and Age(11-17)

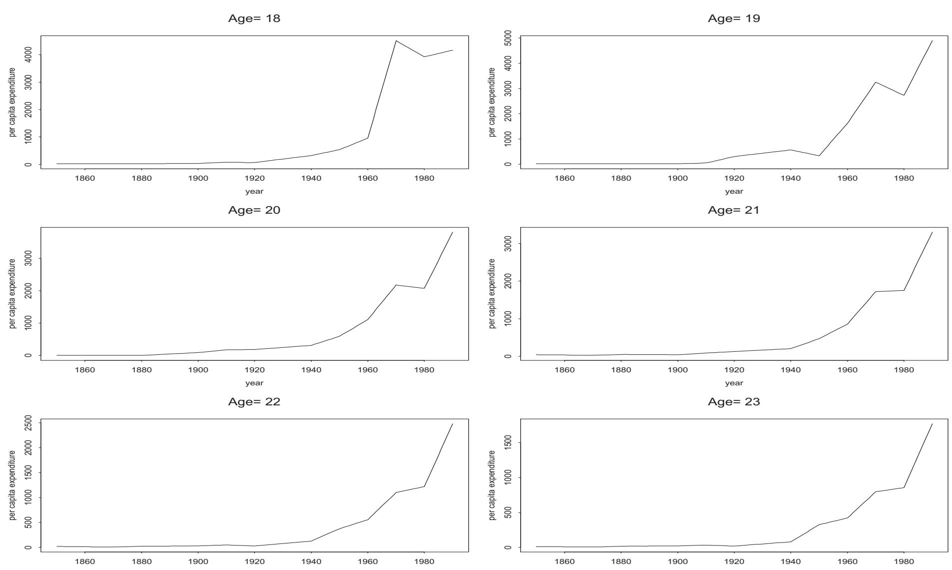

Figure 3: Per Capiture Public Education Expenditure by Time and Age(18-23)

2. The educational tax mainly comes from the property tax, which is proportional to the current value of the property.

3. There are three types of people paying the property taxes: home owners, renters and landlords. About 70% of taxes on rental properties are paid by renters in the form of higher rents. The age profile of homes owned by landlords, who paid the other 30%, is the same shape as the age profile of owned-home values.

4. In the census, the value of the property is reported by respondents who own their own home. Renters report their average monthly rent which we assume is proportional to the value of the property. We use census data from 1940-1990 to derive the age profile of home value for heads who own their homes and the age profile of monthly rent for heads who pay rent.

5. Since data from BEA yields an accurate estimate of the aggregate value of residential housing by tenure: owned and rented, we use this data to adjust the levels of the two age profiles derived by the census.

The detailed procedures are as follows:

2.2 From the Census:

1. Derive the home value for heads who own their homes by ages, h(a,t), for censuses from 1940 to 1990. We then use interpolation for intervening years. For pre-1940 years, we use 1940's age profile.

2. Derive the housing value for heads who rent their homes by age, i(a,t) for census from 1940 to 1990. We use only those who pay rent and derive the approximate housing value by multiplying teh annual rent by 7.[3] Also, we use linear interpolation for intervening years and 1940's data for previous years.

3. Derive the number of heads who own their home by age, o(a,t), for censuses from 1900 to 1990 and we interpolate for the intervening years.

4. Derive population counts by age, p(a,t), for censuses from 1900 to 1990, which is also available from IPUMS. We interpolate for the intervening years.

2.3 From the BEA

We have the aggregate value of owned homes, $\mathbf{M(t)}$, and of rented homes, $\mathbf{N(t)}$, 1925-1990 from the BEA.

[3]This r̈ule of 7ïs frequently used in real estate valuation. We do not need to be concerned about its accuracy, since we are more interested in the age-shape than the level. Thus, this r̈ule of 7g̈ives us a ballpark estimate of housing occupied by renters to compare to the BEA.

2.4 Adjustment and Datacheck

We then adjust the level of the census profiles: h(a,t) and i(a,t) so as to match the aggregate values reported by the BEA:

$$\mathbf{h}\prime(\mathbf{a},\mathbf{t}) = h(a,t)\frac{\mathbf{M}(t)}{\sum h(a,t)o(a,t)} \tag{4}$$

$$\mathbf{i}\prime(\mathbf{a},\mathbf{t}) = i(a,t)\frac{\mathbf{N}(t)}{\sum i(a,t)r(a,t)} \tag{5}$$

In the above equations, $\frac{\mathbf{M}(t)}{\sum h(a,t)o(a,t)}$ and $\frac{\mathbf{N}(t)}{\sum i(a,t)r(a,t)}$ are two adjustment factors for housing value and rental value. We plot both adjustment factors by year on the same graph(see Figure 4). There is no strong time trend and the factors are overall not far from 1.0. So we will use these adjustment factors in later calculations.

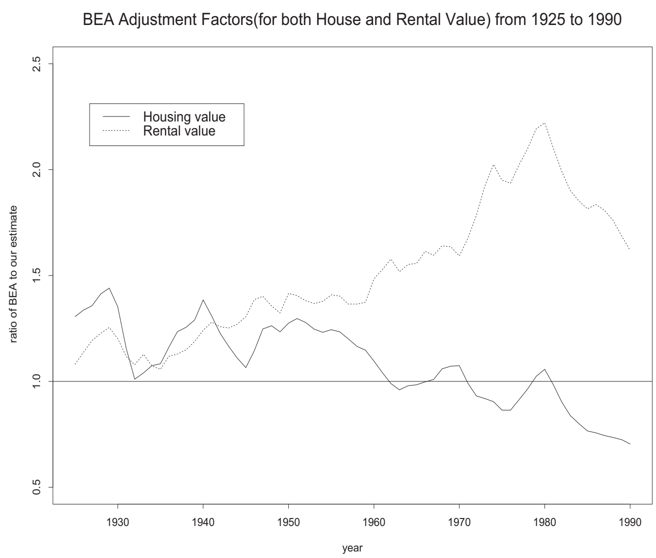

Figure 4: BEA Adjustment Factors over time

2.5 Assigning incidence of property tax and calculating per-capita values:

1. As stated before,although landlords are billed for the property tax due on the rental properties they own, we assume that 70% of this is passed on to renters in the form of higher taxes. We also assume that the age profile of landlords is identical to the age profile of homeowners. W then create two new age profiles to reflect the tax incidence:

(a) An age profile of property for homeowners and landlords:

$$\mathbf{j}(\mathbf{a},\mathbf{t}) = h(a,t)\frac{\mathbf{M}(t) + 0.3\mathbf{N}(t)}{\sum h(a,t)o(a,t)} \tag{6}$$

(b) An age profile of rental property value for renters:

$$\mathbf{k}(\mathbf{a},\mathbf{t}) = i(a,t)\frac{0.7\mathbf{N}(t)}{\sum i(a,t)r(a,t)} \tag{7}$$

2. Next, we combine these two age profiles to create a per-capita age profile of housing value for the years 1925-1990:

$$\mathbf{l}(\mathbf{a},\mathbf{t}) = \frac{\mathbf{j}(\mathbf{a},\mathbf{t})o(a,t) + \mathbf{k}(\mathbf{a},\mathbf{t})r(a,t)}{p(a,t)}, \forall t \in [1925, 1990] \tag{8}$$

Also, for years beyond 1990 and before 1925, we just apply the boundary age profiles, namely 1925's and 1990's age shape of taxation.The **l(a,t)** in the above equation then defines the age shape of taxation before adjustment by eduation expenditure.

3. Since the total educational expenditures equal the total property tax paid, the final tax profile is obtained by multiplying an adjustment factor vector over time:

$$\mathbf{L}(\mathbf{a},\mathbf{t}) = \mathbf{l}(\mathbf{a},\mathbf{t})\frac{\sum_a \mathbf{C}(\mathbf{a},\mathbf{t})p(a,t)}{\sum_a \mathbf{l}(\mathbf{a},\mathbf{t})p(a,t)} \tag{9}$$

And this **L(a,t)** is the final matrix of age shape of taxation for ages 0 to 90, years 1850 to 2090.

3 Survival Weighted, Discounted NPV

1. Once we have the two main matrices(**C(a,t)** & **L(a,t)**), we can trace the expenditure and taxes of any birth cohort by looking at the diagonal values of these period matrices. Let $\mathbf{C}^*(\mathbf{a},\mathbf{t})$ denote the cohort matrix of education expenditure, and $\mathbf{L}^*(\mathbf{a},\mathbf{t})$ the cohort matrix of education taxes, where a is the age of the cohort member who is born at year t, and we have:

$$\mathbf{C}^*(\mathbf{a},\mathbf{t}) = \mathbf{C}(\mathbf{a},\mathbf{t}+\mathbf{a}), \mathbf{L}^*(\mathbf{a},\mathbf{t}) = \mathbf{L}(\mathbf{a},\mathbf{t}+\mathbf{a}) \tag{10}$$

Note that $\mathbf{C}^*(\mathbf{a},\mathbf{t})$ is a 36×151 matrix (with ages 0 to 35 and birth year from 1850 to 2000) and $\mathbf{L}^*(\mathbf{a},\mathbf{t})$ is of dimension 91×151.

2. (a) In order to calculate the NPV, we need to apply survival weights tothese streams of benefits and costs. We obtain cohort survival curves (lx curves) from 1900 onward from the Social Security mortality data posted at the *Berkeley Mortality Database*.[4] For earlier

[4]see http://www.demog.berkeley.edu/wilmoth/mortality/states.html

years, we have another historical source provided by Michael Haines. It has period mortality rates between 1850 and 1900. We did the interpolation by single year and age on the log scale and thus get a matrix of period mortality rates by every year and age. From there, we derive cohort life tables, especially lx values for cohorts born between 1850 and 2000 using demographic formulas. At the end of this step, we have a matrix of survival rates for each cohort, denoted by **S(a,t)**, where a is the age of the cohort member born in year t.

(b) Discounting enters the NPV calculation too. We use two discount values: constant over time at 3% and historical time series of the real interest rates. For the historical time series of interest rates, we have from year 1890 and 1997. We then average the closest 10 years' values and apply them to years before 1890 and years beyond 1997(the whole series is shown in Figure 5). Then we end up with a vector/matrix of discount factors(**D(a,t)**:

$$\mathbf{D}(\mathbf{a},\mathbf{t}) = exp(-0.03 \times a), with flat discounting \quad (11)$$

For discounting with historical interest rates, we have a time vector of historical real interest rates from 1850 to 2090, denoted by **H(t)**:

$$\mathbf{D}(\mathbf{a},\mathbf{t}) = 1, if a = 0 \quad (12)$$

$$= exp(-\sum_{a=0}^{90} \mathbf{H}(\mathbf{t}+\mathbf{a})), otherwise \quad (13)$$

3. Finally, we calculate the NPV of education for birth cohorts using the survival-weighted, discounted present value of benefits minus costs.

$$\mathbf{NPV}(\mathbf{t}) = \sum_a \mathbf{C}^*(\mathbf{a},\mathbf{t})\mathbf{S}(\mathbf{a},\mathbf{t})\mathbf{D}(\mathbf{a},\mathbf{t}) - \sum_a \mathbf{L}^*(\mathbf{a},\mathbf{t})\mathbf{S}(\mathbf{a},\mathbf{t})\mathbf{D}(\mathbf{a},\mathbf{t}) \quad (14)$$

The three different versions of the NPV(undiscounted, discounted with 3%, and with historical interest rates) are shown in Fig 6, 7 and 8.

4 Other important calculations

1. In addition to the NPV calculation, we also compute the NPV as a percentage of lifetime earnings from 1870 to 2000. We made a big assumption that the age profile of earnings in any year was the same shape as the 1990 age profile. For 1990's age profile, we have the aggregate data of earnings for the whole population. We also assume its shape is the same as that in year 1999 which is available and adjust its level with the aggregate data. After obtaining year 1990's earning profile by age, we use GDP per capita as adjustment tool to raise or lower the level for other years. In addition, the calculation of lifetime earnings is also discounted and survival

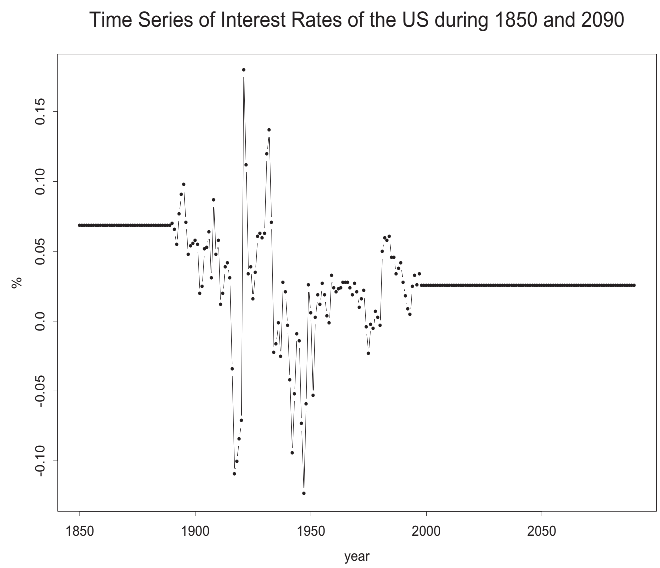

Figure 5: Historical Real Interest Rates

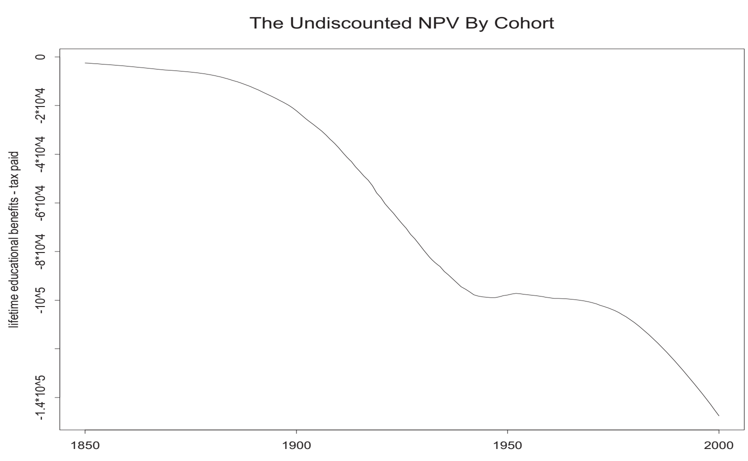

Figure 6: NPV

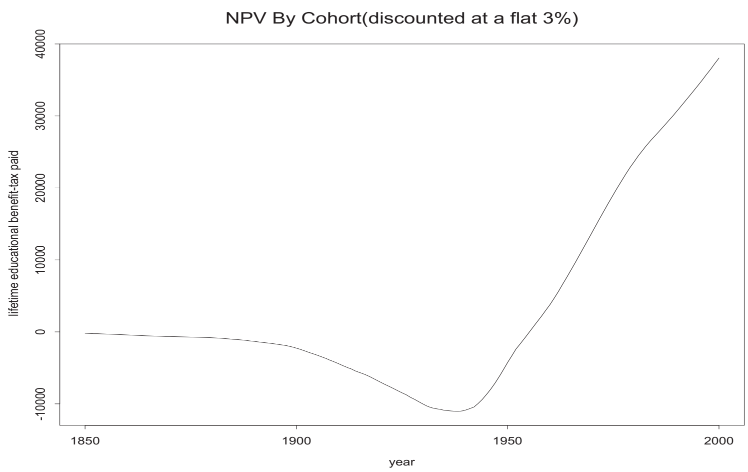

Figure 7: NPV(3%)

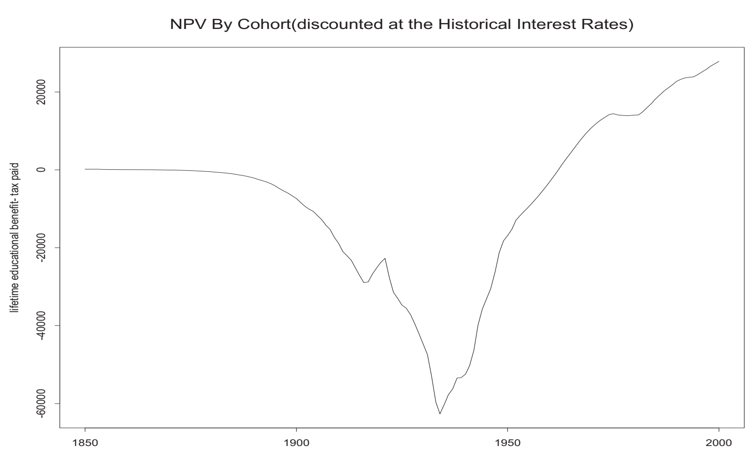

Figure 8: NPV(historical interest rates)

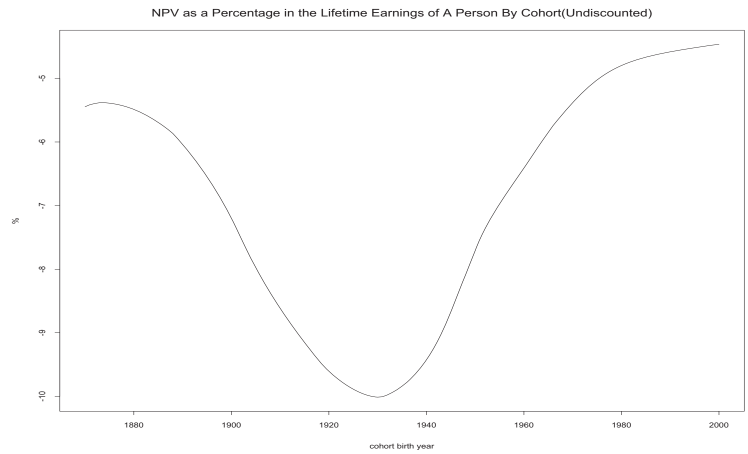

Figure 9: NPV as a Percentage of Lifetime Earnings(undiscounted)

weighted, with predictions made for period data beyond 1990 based on the 1.8% labor productivity growth rate as well.
The resulting NPV as a percentage of lifetime earnings in three versions are in Fig 9, 10 and 11. The plots of lifetime earnings by cohort are also attached in Fig 12, 13 and 14.

2. It is also interesting to decompose the NPV into its two PV components: the present value of educational benefits and the present value of educational costs. They are both survival weighted and discounted, as shown in Fig 15, 16 and 17.

3. We also calculate the mean age of receiving education benefits and paying education taxes for each cohort. Let $\mathbf{EB(t)}$ denotes the mean age of receiving benefits and $\mathbf{ET(t)}$ be the mean age of paying taxes:

$$\mathbf{EB}(\mathbf{t}) = \frac{\sum_{a=0}^{35} \mathbf{C}^*(\mathbf{a},\mathbf{t})\mathbf{S}(\mathbf{a},\mathbf{t})\mathbf{D}(\mathbf{a},\mathbf{t}) \times a}{\sum_{a=0}^{35} \mathbf{C}^*(\mathbf{a},\mathbf{t})\mathbf{S}(\mathbf{a},\mathbf{t})\mathbf{D}(\mathbf{a},\mathbf{t})} \tag{15}$$

$$\mathbf{ET}(\mathbf{t}) = \frac{\sum_{a=0}^{90} \mathbf{L}^*(\mathbf{a},\mathbf{t})\mathbf{S}(\mathbf{a},\mathbf{t})\mathbf{D}(\mathbf{a},\mathbf{t}) \times a}{\sum_{a=0}^{90} \mathbf{L}^*(\mathbf{a},\mathbf{t})\mathbf{S}(\mathbf{a},\mathbf{t})\mathbf{D}(\mathbf{a},\mathbf{t})} \tag{16}$$

The results are in Figure 18 and 19. The mean age of receiving educational benefits (i.e. going to school) ranges from 13 to 16, with the lowest value in 1850 and highest in 1949. After 1950, the mean age of receiving educational beneifts started to decline, which is quite counter intuitive. By examining the data, we found that is because the college expenditure increased much more rapidly around 1950 relative to the spending on elementary and secondary school, which lifted the mean age of receiving

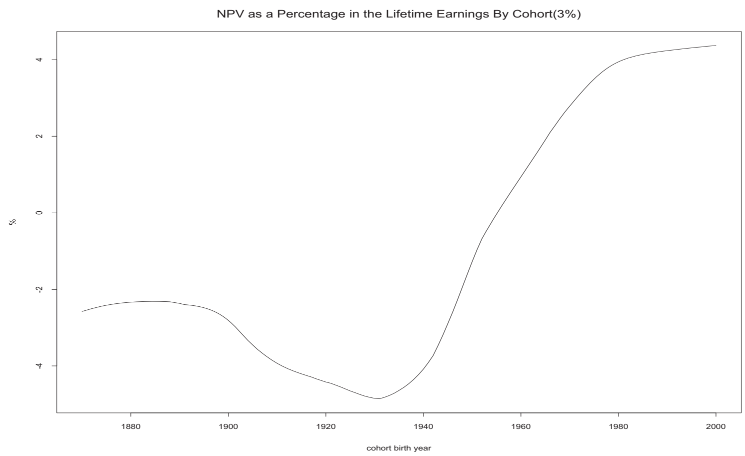

Figure 10: NPV as a Percentage of Lifetime Earnings

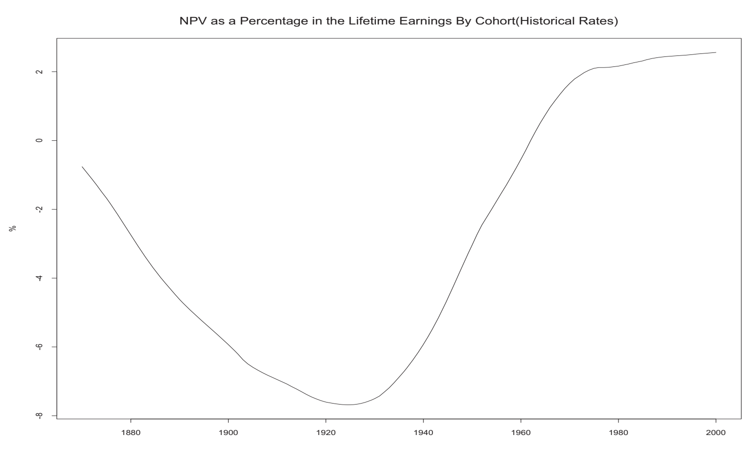

Figure 11: NPV as a Percentage of Lifetime Earnings(historical rates)

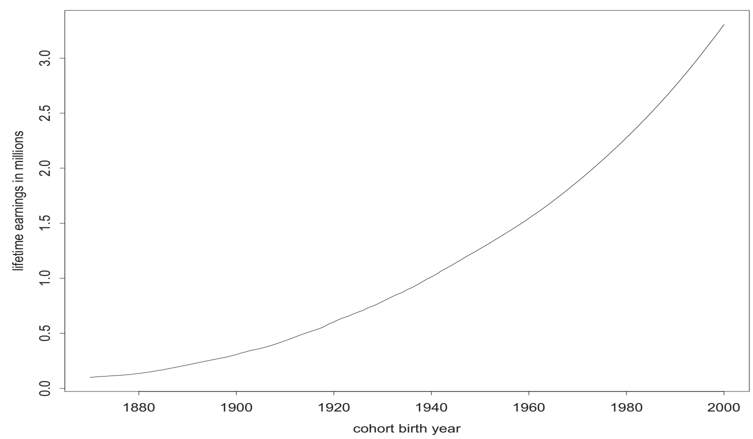

Figure 12: Lifetime Earnings by Cohort(undiscounted)

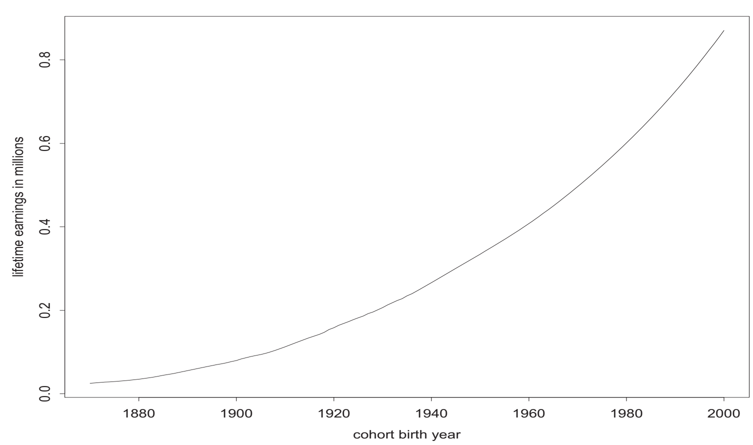

Figure 13: Lifetime Earnings by Cohort(3 %)

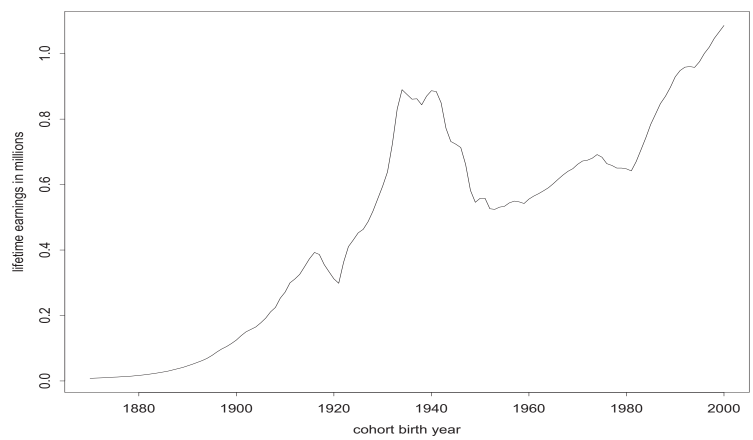

Figure 14: Lifetime Earnings by Cohort(historical interest rates)

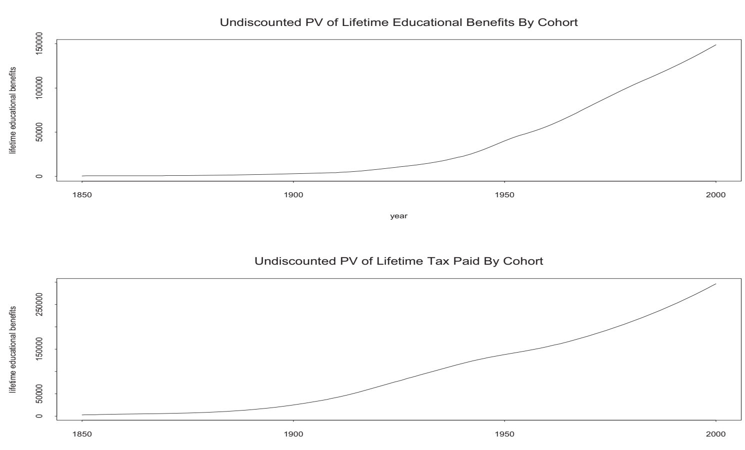

Figure 15: Present Values of Educational Benefits and Costs

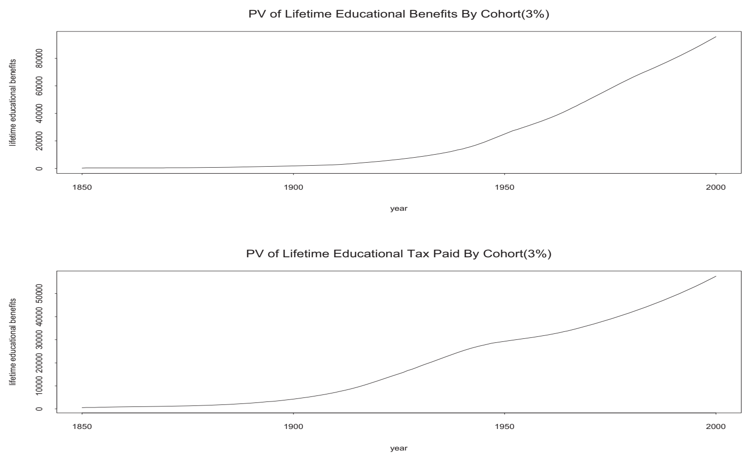

Figure 16: Present Values of Educational Benefits and Costs(3%)

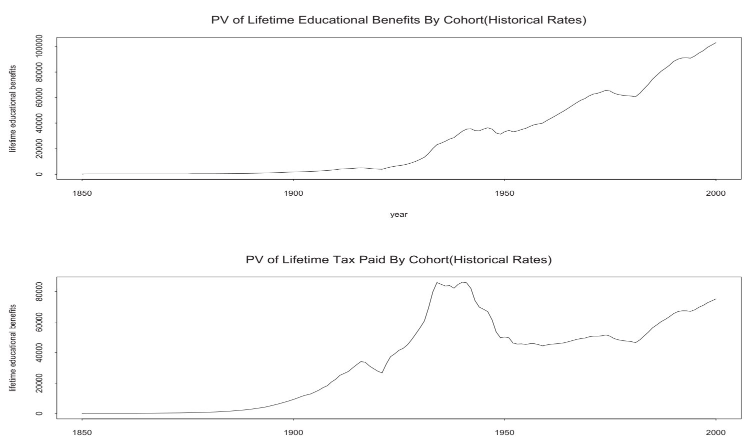

Figure 17: Present Values of Educational Benefits and Costs(historical rates)

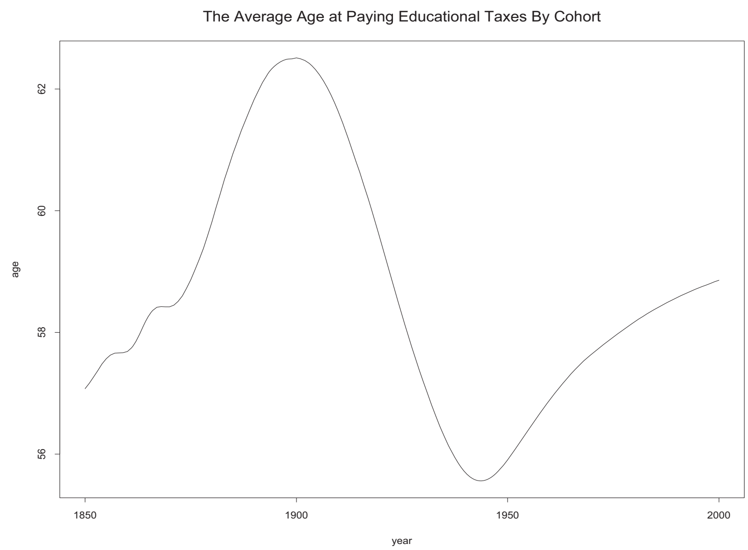

Figure 18: Mean Age of Paying Taxes by Cohort

educational benefits. However, after 1950,college spending remained quite stable while the elementary and secondary education expenditure continued to increase. As a result, the gap between per capita expenditure before age 18 and after 18 diminished and it pulled down the mean age relative to the peak in 1949.
On the other hand, the mean age of paying educational taxes ranges from 55 to 63, occurring in 1900 and 1944, respectively. The old ages seems suspicious at first, so as a data check, we calculated the period mean age of paying property tax in 1992 from the Current Population Survey. The period mean age is 57.5 in 1992, which is consistent with our period estimates. In addition, the cohort result is bigger than the period estimate in years after 1990. This is because though after 1990 we assume a constant age shape and use 1.8% growth rate as well as the 3% discounted rate, the money value of tax payments in older ages in later periods are still big enough to offset the -1.2growth. In other words, even when discounted and survival weighted, the tax payments for a cohort member still increases with age. So by comparing the 2000 period tax profile and the 2000 cohort tax profile, we see that the period's tax profile declines at an earlier age than the cohort's tax profile. So that leads to a higher cohort mean age of paying taxes.

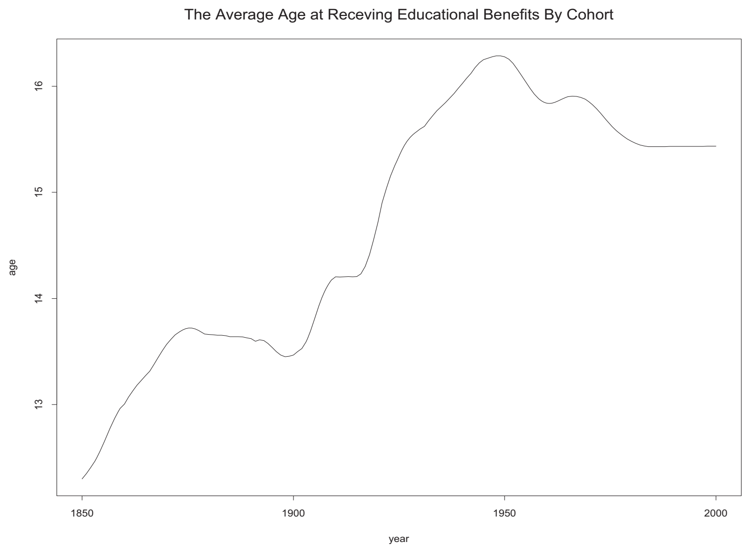

Figure 19: Mean Age of Paying Taxes by Cohort

Part IV
Results and Discussion

The biggest finding from this calculation is that people who were born around 1940 have the most negative net present value of educational benefits minus educational taxes. In other words, they are the biggest losers in terms of paying more educational taxes (net receiving its benefits) than other cohorts in lifetime. The NPV reaches minimum at birth year 1938 with value well over $10,000 (constant 1999 dollar) in magnitude(Figure 6, 7 and 8). If we look at the NPV as a percentage of lifetime earnings, the same people lose about 4% of their lifetime earnings for paying educational taxes (net benefits)(see Figure 9).
There could be two reasons behind this:

1. Population Growth Effect (or the baby boom phenomenon).
   For cohorts born around 1940, they have to burden the educational costs for the baby boomers who entered school at the same time they started paying taxes (particularly property taxes). Since the bulk of the educational expenditures come from college spending, and when baby boomers went to college (at around year 1970), the late-1930s-birth-cohorts were in their early 30s and started to form families, buying houses and paying property taxes. Because of the unusual large number of the babyboom generation, the late-1930s-cohorts are definitely paying more (property) taxes than other cohorts, as the babyboom generation was both preceded

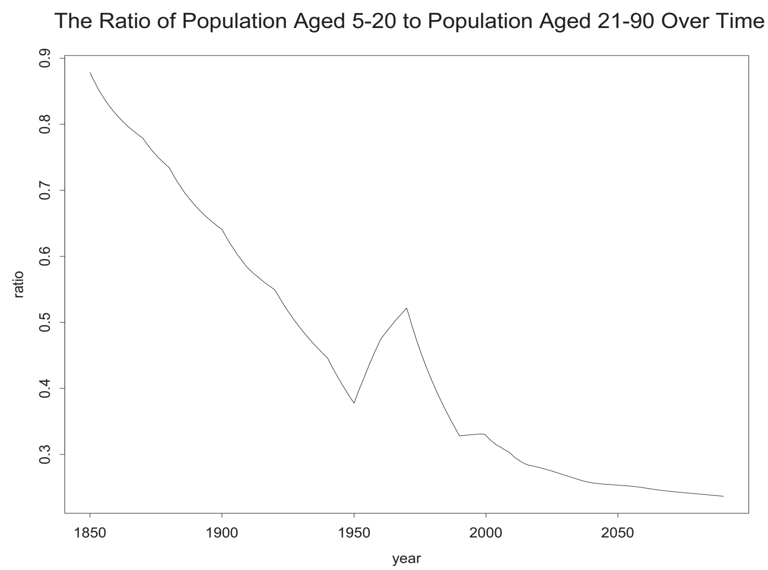

Figure 20: Population Ratio

and followed by a much smaller birth cohort. This effect is also illustrated by a graph of the population ratio of the age 5-20 population and the age 21-90 population (Figure 20). There is a decline trend of the ratio over time, except for the period of the babyboom, where the ratio jumped tremendously. Another plot(Figure 21) shows the growth rate of the (normalized) period tax paying population over time[5] It is clear from the plot that the growth rate is particularly low for the period 1970s (equivalent to cohort born in the 1930s-1940s) and it rises again for the period 1990s (or the 1950 and after birth cohorts). In addition, from Fig 22, we see that total expenditure growth rate peaks in the 1970, which corresponds to the babyboom birth cohort, and it is the lowest around year 1950 that refers to the cohort born around 1930s to 1940s.

2. Education Expenditure Growth effect(on the per-capita basis).
   As stated before, college spending consitītues a big part of education expenditure. Therefore, a dramatic increase in college enrollment for some cohort, and thus an increase in per-capita public education expenditure will cause the present value of educational benefits for the very cohort to increase and also raises the present value of educational costs for the cohort who are paying for it. By looking at Figure 1, we find that the

[5] Normalization means tax-weighted population. We first calculate the tax payment at each age as a percentage of all ages in that year. Then we average it over the whole time series for each age, with which we time the population profile to get the normalized tax-paying population.

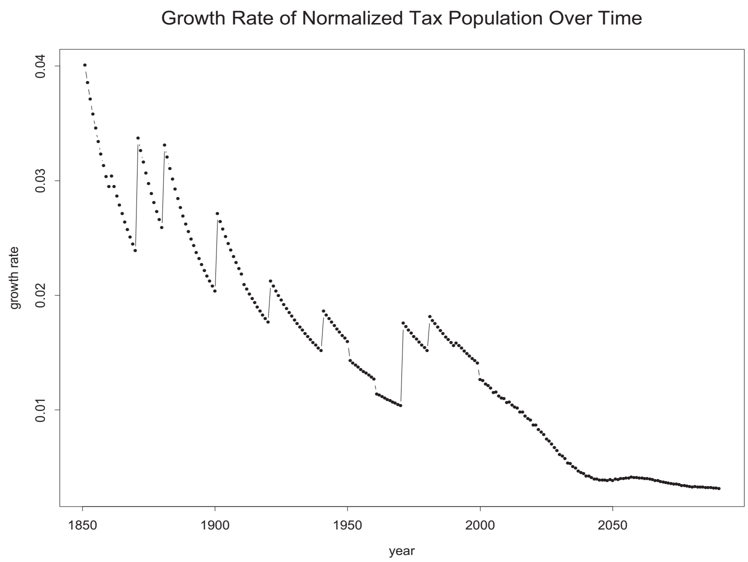

Figure 21: Normalized Tax-Paying Population

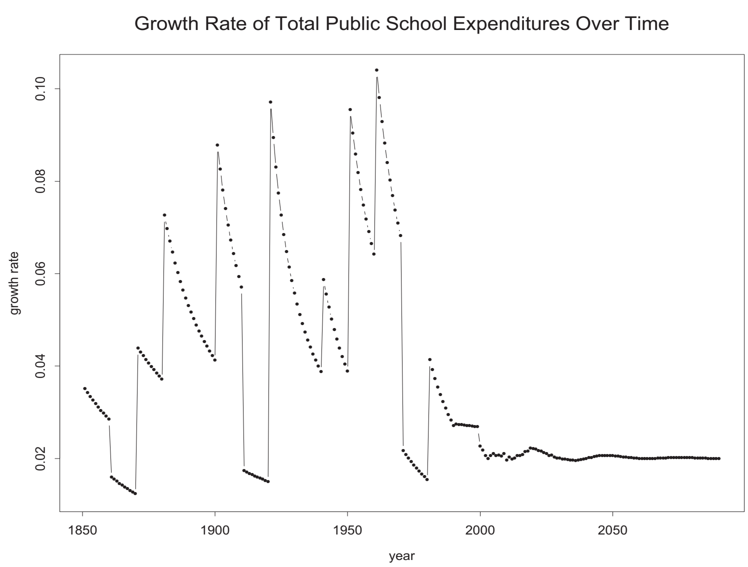

Figure 22: Total Public Education Expenditure Growth Rate Over Time

babyboomer's per-capita college expenditure ($>$ age 18) soared in the middle 1970s, which deviates a great deal from the gradually increasingly trend. It is particularly obvious at age 18, when the per-capita public educational expenditure has a smooth trend over time, except in the 1970s when there was a dramatic increase. Therefore, in addition to the increasing number of students that share the benefits, the per-capita expenditure also rose greatly most probably because of larger college enrollment of the babyboomers.

Those effects can also been interpreted from observing the mean age of receiving benefits and paying taxes. From Figure 19, it is easy to see that the 1950 birth cohort(the babyboomers) has the highest mean age of receiving educational benefits, which implies a big educational expenditure burden for the concurrent tax paying cohorts.

In early years, there are various forces which worked in the opposite directions. Population was increasing at a higher rate but the college enrollment was low and per-capita expenditure was not so high. So the effects were not as prominent. For years after 1950, the NPV keeps increasing, partly because the population is growing at slower rates so that each generation/cohort is paying educational taxes for a smaller generation. In addition, the enrollment (especially in higher education) increases more slowly or levels off gradually. Therefore, though per-student expenditure still increases, the effect on per-capita expenditure and on educational costs imposed on the tax-paying cohorts get smaller.

Another informative plot is that of the mean enrollment ratio between cohorts born at time t and cohorts born 35 years later(see Figure 23). The mean enrollment is the arithmetic average of enrollment rate for a cohort member in his lifetime (i.e. the average of enrollment from age 5 to 35). Although this series only lasts through the 1920s[6], we see that the ratio is increasing till the birth year 1880 and reaches the minimum around year 1905. It means that the 1940 birth cohort has a lower lifetime enrollment relative to the cohort that was paying for its education. This is in agreement with our finding that the NPV for the 1940 birth cohort is the lowest.

Part V
Comparison

In addition, we also include the NPV calculations from another Social Security project for comparison. The data used in the study is for the NPV of OASI (Oldage and survivor's insurance). That is, they do not include the DI (disability

[6]Since we calculate on a cohort basis, we can only have the average enrollment for cohorts born in the 1960s, and then we need to get the ratio between cohorts that are 35 years apart. That is why we end up having only around 75 years of this ratio.

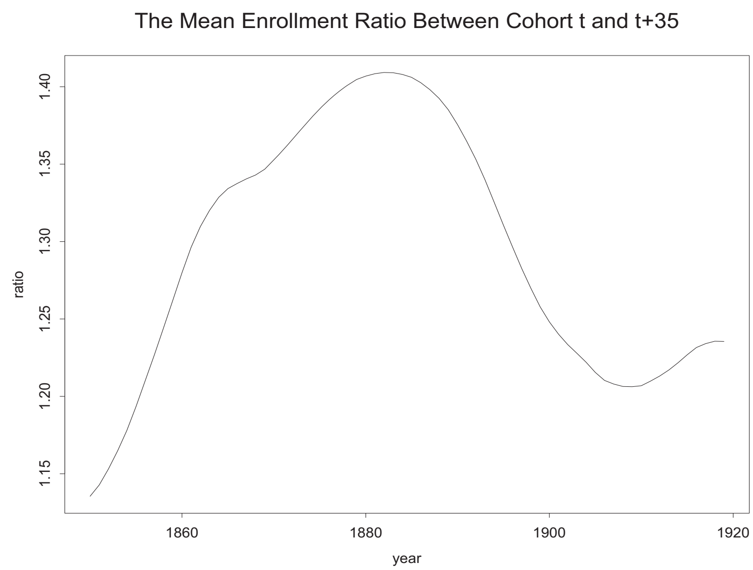

Figure 23: Mean Enrollment Ratio between Cohorts t and t+35

insurance) or HI (health insurance) part of social security. They are discounted at 3%. For the forecasts it is assumed that beginning in 2033 (the year the OASI trust funds were projected to become bankrupt), taxes are adjusted in every subsequent year so that total annual OASI benefits are paid. (That is, it is assumed the OASI system becomes "PAYGO" in 2033).[7] Fig 24 shows the present value of the benefits and costs for birth cohorts from 1880 to 2060. Fig 25 is the NPV, Fig 26 is the NPV as a percentage of lifetime earnings.

For comparison, we also plot the NPV and NPV as a percentage of lifetime earnings from both studies on the same graph(Fig 27 and 28). It is interesting finding that early cohorts paid more taxes than the amount they received for education but enjoyed a windfall gain from the social security system. The gain from the social security system gradually declines while the net benefits from public education decreases to the lowest in 1938 before it starts to rise. The current trend is that people are receiving more in public education but are also paying more in social socurity. Overall, the the late 1880s and 1980 birth cohorts are the winner in both systems, and the 1930-1940 birth cohorts suffer the biggest losses from the two systems(Fig 29 and 30 shows the sum of the NPV(in real terms and percentage of lifetime earnings from the two systems).

[7]The data were used in the paper: Ronald Lee, Andrew Mason, and Tim Miller (2001) Saving, Wealth, and Population, Nancy Birdsall, Allen C. Kelley and Steven W. Sinding, eds., Population Does Matter: Demography, Poverty, and Economic Growth. Oxford University Press, in press.

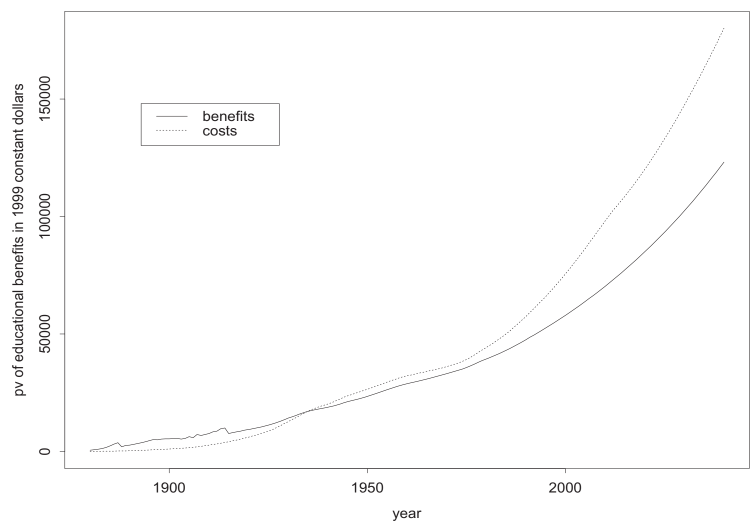

Figure 24: Present Value by Cohort in Social Security Studies

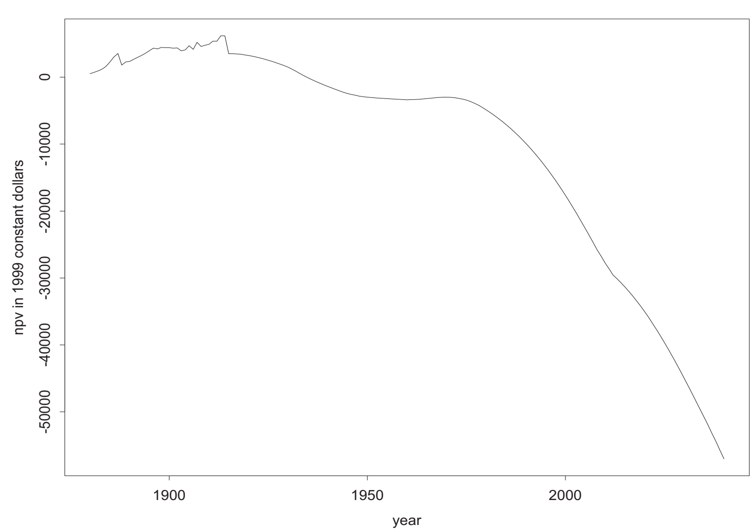

Figure 25: NPV from Social Security Studies

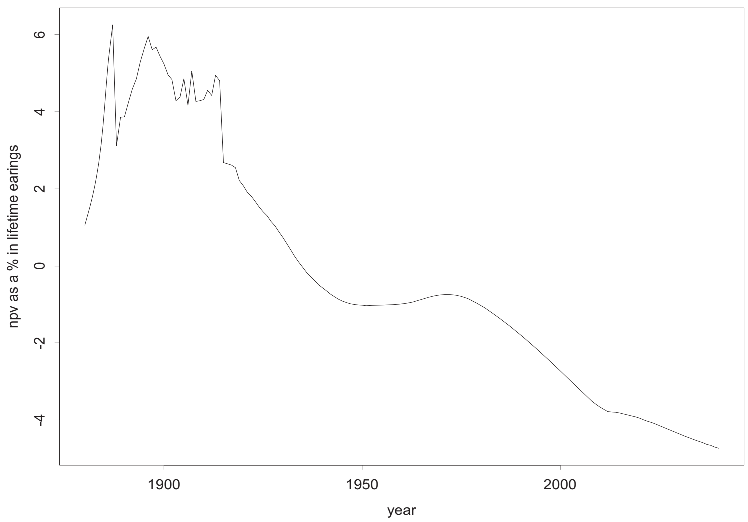

Figure 26: NPV as Percentage of Lifetime Earnigns from Social Security Studies

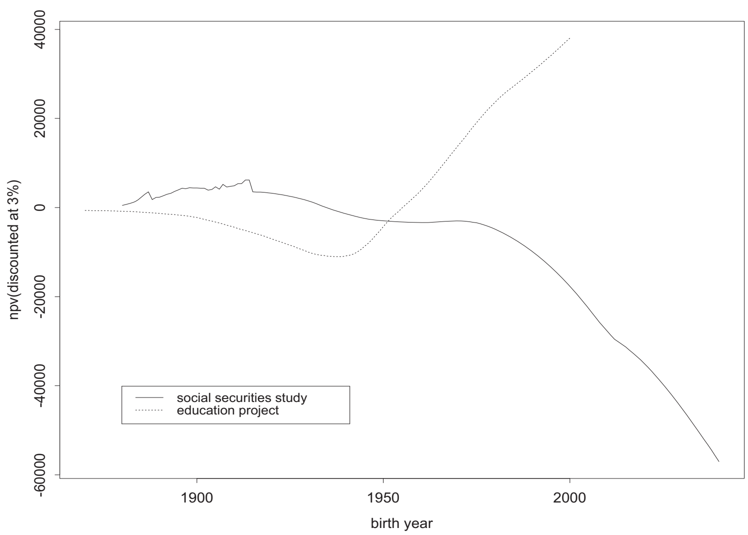

Figure 27: Comparison of NPV(Discounted at 3%)

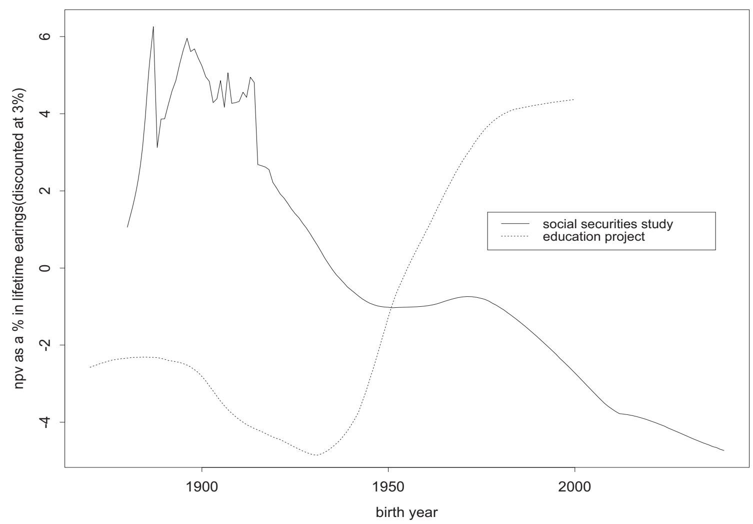

Figure 28: Comparison of NPV as a % of Lifetime Earnings (Discounted at 3%)

Figure 29: The Sum of two NPV (Discounted at 3%)

Figure 30: The Sum of two NPV as a % of Lifetime Earnings (Discounted at 3%)