2010

2010-4 Human Capital Prices, Productivity and Growth

Audra J. Bowlus
University of Western Ontario, abowlus@uwo.ca

Chris Robinson

Human Capital Prices, Productivity and Growth

by

Audra J. Bowlus and Chris Robinson

Working Paper # 2010-4 **November 2010**

CIBC Working Paper Series

Department of Economics
Social Science Centre
The University of Western Ontario
London, Ontario, N6A 5C2
Canada

Human Capital Prices, Productivity and Growth *

Audra J. Bowlus
University of Western Ontario†

Chris Robinson
University of Western Ontario ‡

November 2010

Abstract

Separate identification of the price and quantity of human capital has important implications for understanding key issues in economics. Price and quantity series are derived for four education levels. The price series are highly correlated and they exhibit a strong secular trend. Three resulting implications are explored: the rising college premium is found to be driven more by relative quantity than relative price changes, life-cycle wage profiles are readily interpretable as reflecting optimal human capital investment paths using the estimated price series, and adjusting the labor input for quality increases dramatically reduces the contribution of MFP to growth.

*The authors wish to thank Lance Lochner and Todd Stinebrickner for helpful comments and discussion. We also thank participants in seminars at the University of Virginia, the University of Guelph, McMaster University, the University of British Columbia, Wlifrid Laurier University and conference sessions at the 2006 CEA Annual Meetings, the first annual UM/MSU/UWO Labor Economics Day, and the CIBC Human Capital, Productivity and the Labour Market conference. This work was supported by the CIBC Human Capital and Productivity Centre and the Canadian Social Sciences and Humanities Research Council.

†Department of Economics, University of Western Ontario, London, ON N6A 5C2, Canada, e-mail: abowlus@uwo.ca.

‡Department of Economics, University of Western Ontario, London, ON N6A 5C2, Canada, e-mail: robinson@uwo.ca.

1 Introduction

The flow from human capital is by far the most important input in the world economy. The estimated share of labor in the U.S. and most of the OECD countries, for example, is about two thirds. There is quite general agreement that human capital plays a significant role in the determination of living standards. Human capital theory has been the basis of a huge literature studying the determination of earnings and inequality.[1] The recent rise in the skill premium and inequality in the U.S. has stimulated a large literature, based on the human capital framework, to provide an explanation.[2] At the same time, it has exposed a continuing weakness of human capital theory due to an inherent under-identification problem. This is a problem for many standard microeconomic based analysis of earnings patterns, but it is also a problem for macroeconomic growth studies. Payments to human capital may be directly observed, but a payment is a product of a price and a quantity. In general, neither the quantity nor the price of human capital is directly observable. Thus, when payment (wage) differences are observed between any observable groups, say by education level or by country, in general it is not possible to distinguish whether the difference is due to a difference in prices or a difference in quantities, or some combination of the two. This has important implications for our understanding of many key issues in labor economics and macroeconomics, including the rising college premium, rising inequality, sources of growth and life-cycle productivity profiles. With a small number of exceptions, the vast literature on human capital has ignored this identification problem.[3]

In most cases implicit identification assumptions are made, usually representing one of two extremes: either constant quantities or constant prices over time. Assessing the contribution of human capital to output, living standards and growth is hampered by serious conceptual and measurement problems due to this identification issue. It is well recognized that the quantity of the labor input cannot simply be measured by total hours.[4] The implicit identification strategy typically adopted to

[1]Seminal works include Becker (1964), Ben-Porath (1967), and Mincer (1974).

[2]See, for example, Katz and Murphy (1992), Card and Lemieux (2001), and the survey by Katz and Autor (1999), for traditional labor economics approaches, and Krusell *et al.*(2000), Guvenen and Kuruscu (2007) and Huggett, Ventura and Yaron (2006) in the macroeconomics literature

[3]The relatively few exceptions include Heckman, Lochner and Taber (1998), Weiss and Lillard (1978), and Huggett, Ventura and Yaron (2006).

[4]The importance of this problem was recently emphasized by Solow (2001, p.174): "alternative ways of measuring human capital can make a big-time difference in the plausible interpretation of economic growth, so it is really important to come to some scientific agreement on the best way to deal with human capital as in input (and as an output, don't forget) and then to implement it."

construct an alternative to hours for a given country is that quantities within an observable "type" of labor are constant and that all wage changes are due to changes in prices. The college premium literature provides a similar example where relative payments are typically taken as relative prices which implicitly assumes that relative quantities are constant.

The early literature on human capital in a life-cycle context studied the implications of optimal investment in human capital for log wage profiles by age or experience. Following influential papers by Ben-Porath (1967), Heckman (1976), Rosen (1976) and others, human capital theory became the dominant framework for analyzing life-cycle earnings. This framework has spread from its base in labor economics and is increasingly used in the modern macroeconomics literature. Recent examples include Huggett, Ventura and Yaron (2006) and Guvenen and Kuruscu (2007).[5] Empirical examination of experience profiles, based on the human capital framework, has a long tradition going back to Mincer (1974). The Mincer inspired literature has recently been re-examined by Heckman, Lochner and Todd (2002). The large number of papers indicates the continuing importance of this topic for economists in understanding wage patterns. This life-cycle literature provides a contrasting example of the standard identification approach, representing the other extreme, where the life-cycle profile of payments is typically assumed to be the same as the life-cycle profile of the quantity of human capital which implicitly assumes that the price is constant over the life-cycle.

In this paper we address the identification problem directly and construct price series for four possible types of human capital associated with commonly used education groups: high school dropouts, high school graduates, some college and college graduates. There is no simple solution to the problem. Some assumptions have to be made. Our assumptions are discussed in detail so that they may be contrasted with the implicit, and often in our view, extreme assumptions generally made in the literature. The price series that are constructed for these four education groups turn out to be surprisingly highly correlated over the period 1963 to 2008 despite their large differences in skill level. The series also exhibit patterns that deviate substantially from wages, implying that wages are not good proxies for prices and that quantities of human capital associated with a given observable education type change over time. Both of these results have major implications for understanding the evolution of wages, wage premia and the human capital input in the aggregate production function (and hence, total factor productivity).

[5]A separate but related development in the literature has been the estimation or calibration of general equilibrium models of human capital accumulation over the life cycle. Examples include Heckman, Lochner and Taber (1998), Imai and Keane (2004), Lee (2005) and Hansen and Imrohoroglu (2007).

In Section 2 the basic identification issue is discussed. This is not a problem unique to human capital. The possible change in quantities of human capital associated with a given observable education type over time or across countries is identical to the problem of an unobserved change in quality over time in a product. A prominent example that has received much attention in the macroeconomic literature is the identification of an appropriate price series for various forms of physical capital inputs, particularly the input from computers. It is clear that conventional methods of constructing a measure of inputs over time in the case of computers fails very badly. Measures based on physical numbers, price or total value all drastically underestimate the computer input. The identification issue has received a lot of attention. Unfortunately, due to observability issues, the solution for human capital is more difficult than for many types of physical capital.

Section 3 details our identification approach and implements it on data from the U.S. March Current Population Surveys (MCPS) covering earnings years 1963 to 2008. A flat spot method is used to construct separate price series for the four human capital types. This method exploits variation over short intervals of time towards the end of the life-cycle in payments to the same cohort of workers. A methodology is developed for choosing the flat spot that differs from Heckman, Lochner and Taber (1998). The price series all show substantial movement over the 1963 to 2008 period and the patterns are robust to a number of validity checks and sensitivity analysis. Perhaps the most surprising result is a very high correlation between the series from the lowest education group (high school dropouts) to the highest (college graduates). All of the series exhibit an increase in the price from 1963 to the mid-1970s followed by a substantial decline through the 1980s and 1990s that is interrupted by plateaus or recoveries coming out of the recessions of the early 1980s and the early 1990s. Finally, all four prices are relatively flat after 2001.

The price series of Section 3 have many important implications for both microeconomic and macroeconomic issues. First, since the series are highly correlated, the movement in relative prices is quite modest implying that a substantial part of the variation in the college premium is due to variation in the relative quantities of human capital associated with each observed education group. Second, since the price varies over the life-cycle, the path of a cohort's wages or earnings does not identify the path of human capital. Third, the high correlation also implies that a homogeneous model may be a good approximation for macroeconomic models dealing with secular trends. This has the great advantage of a simplified and easily interpretable labor input which may be constructed by dividing total payments to labor by the estimated price series. The use of this measure of the labor input suggests a very different path for total factor productivity (TFP) than is derived from conventional aggregate labor input measures. Each of these implications is explored in turn.

Section 4 re-examines the college premium, and decomposes the change in relative wages into the separate components of relative price and quantity changes. It shows that relative quantity changes are at least as important as relative price changes in explaining the path of the average college premium. In addition, selection effects on relative quantities implied by cohort education patterns can also explain the age patterns, documented in Card and Lemieux (2001). Section 5 shows that the life-cycle pattern of quantities of human capital for each cohort implied by the price series is consistent with human capital theory and, in contrast to the usual constant price assumption, yields readily interpretable cross cohort patterns.

The maintained hypothesis explored in this paper is that college graduates have the same "type" of college graduate human capital over time. In Section 6 we contrast this hypothesis with an alternative skill biased technological change hypothesis under which there was a change in vintage "type", coinciding with the widespread introduction of micro computers, that resulted in an increase in the college premium for the young ("new vintage") college graduates, but not for the ("old vintage") older college graduates, as described in Card and Lemieux (2001). The primary difference between the two hypotheses is their different implications for the time path of vintage effects. Under the hypothesis explored in this paper vintage effects occur through selection effects and technological improvement in human capital production which occur continuously. The alternative hypothesis is that vintage effects come about through changes in vintage types rather than through quantity changes due to selection effects and technological improvement. Thus, for the maintained hypothesis of this paper, the time path of the vintage effects is constrained by the selection effects implied by cohort education patterns combined with technological improvement in the production of human capital. For the alternative hypothesis the pattern is constrained by the timing of the vintage type change. The different constraints result in different predictions for the pattern of vintage effects. Our test results show that the evidence is consistent with our maintained hypothesis and inconsistent with the alternative.

Section 7 uses the estimates of the price series from Section 3 to construct measures of the total labor input and a new TFP series for the U.S. over the period 1975-2003. The results show that conventional quality adjustment to the labor input results in substantial underestimation in the growth of the true labor input, and hence a large overestimate of increases in TFP. Section 8 provides some final discussion and summary.

2 The Price and Quantity of Human Capital: Basic Identification Issues

In standard human capital models with competitive firms the hourly wage is the product of a price and a quantity

$$w_{it} = \lambda_t E_{it}, \tag{1}$$

where E_{it} is the amount of human capital supplied to the firm (number of efficiency units) by worker i in time period t, and λ_t is the rental price paid for renting a single unit of human capital (the price of an efficiency unit). The hourly wage is observed, but its two components are not. This is the fundamental under-identification property of human capital models.

In a homogeneous human capital model there is a single price, λ_t, and wages differ across workers in any given time period because of differences in the amount of (homogeneous) human capital they supply. Over time a worker's wage could change either because of a change in the quantity of efficiency units supplied, or because of a change in the price. Consequently, relative wages between any two observable "types" of workers - say college versus high school graduates - may change over time, but not because of a change in relative prices. All relative wage changes are due to relative changes in the quantity of efficiency units supplied by each type. This is the main consequence of the efficiency units approach in a homogeneous human capital model.

In heterogeneous human capital models, an efficiency units approach is retained within some exogenously defined worker type (e.g. college graduate) but is abandoned across types. With two worker types (e.g. college and non-college) there are two factors and two prices with wages given as follows (suppressing the individual and time subscript for convenience): $w^a = \lambda^a E^a$ and $w^b = \lambda^b E^b$ where λ^a and λ^b are the prices of efficiency units of type a and b, respectively, and E^a and E^b are the number of efficiency units of type a and b supplied by type a and b workers, respectively. Within type, the wage implications are the same as the homogeneous human capital model. For relative wages across types the implications are potentially different. Since there are now two prices, changes in relative wages between the two types reflect changes in relative quantities, E^a/E^b, and changes in relative prices, λ^a/λ^b.

Identification of the prices and quantities of human capital is a difficult problem in both homogeneous and heterogeneous human capital models. In heterogeneous human capital models used in the skill-biased technological change literature, it is implicitly solved by assuming that the quantities of human capital associated with any observed education level at any point in time are the same. This permits the identification of the skill price ratio from the wage ratio. However, this is a very strong assumption. It rules out selection in the choice of who goes on to higher levels of education, changes in optimal life-cycle accumulation of human capital within type, and technological improvement in human capital production functions.[6]

Figure 1 plots completed schooling level for male birth cohorts from 1931 to 1967. It shows that over these cohorts there have been very large changes. For example, the fraction of the birth cohort going to college increased by 50 % between the 1937 and 1946 birth cohorts (from below 20% to over 30 %.) Since there is a strong correlation between measures of ability and the highest level of completed schooling, these large secular changes in completed schooling levels may be expected to have significant selection effects on the average ability, and hence human capital, associated with each observed schooling level.[7] Further, since major quality improvements due to technological change have been found for capital inputs such as computers, it is surprising that they have been generally ruled out for the labor input. We argue that technological improvement in human capital production functions can produce workers of a given observed type, such as education level, that can do more, in the same sense as more recent computers can do more, especially at the upper level as knowledge advances. More recent vintages of physics PhDs, for example, are likely to have received more value added (more advanced knowledge) through the education process, than earlier PhDs.

The identification problem and technological change in human capital production functions was discussed in a related context in Weiss and Lillard (1978) who tried to distinguish between "time" and "vintage" effects in the earnings of scientists. They documented the fact that there was a difference in the life-cycle path of earnings for the more recent vintages of scientists in the period they studied. Since they were unable to employ a separately identified price series of the type derived

[6] If technological change in human capital production is not taken into account, the labor input in aggregate production functions will be underestimated which results in an apparent technological improvement or TFP increase in the product market production functions. Similar mis-attribution can happen with capital mis-measurement. Greenwood, Hercowitz and Krusell (1997) investigate this issue and provide estimates to suggest that the magnitude in the capital input case is important. Estimates in Section 7 below suggest that the magnitude in the labor input case is also very important.

[7] There is a large amount of evidence suggesting a high correlation between observed ability measures and educational attainment. For example, the correlation between highest degree completed and scores from the Armed Forces Qualification Test for individuals in the 1979 National Longitudinal Survey of Youth is 0.58.

in this paper, Weiss and Lillard (1978) were cautious about their interpretation of the vintage effects that they found. Nevertheless, they did conclude that there were substantial vintage effects over the period that they studied. This suggests that the assumption of no change in quantities within observable types may well be a bad one.

The primary maintained hypothesis in this paper is that there are a number of "types" of human capital, as indicated by education group, that are the same over time, allowing for a meaningful definition of changes in quantities within type over time. This is the standard efficiency units assumption within types that is common in the literature. In this framework, the time effects of Weiss and Lillard (1978) can be interpreted as price effects, while the vintage effects can be interpreted as quantity differences across cohorts due to selection effects and technological improvements in human capital production functions, as well as changes in the optimal investment level over the life-cycle by each cohort. This is a parsimonious specification of vintage effects since it maintains a constant number of human capital types and prices. The main objective of the paper is to identify prices and quantities within this framework. However, one interpretation of changes in relative wages by skill is that vintage effects may take the form of introducing new types of human capital (and hence new prices) that are not directly comparable to the old types. This complicates an already difficult identification problem by introducing more types. In general it is difficult to distinguish quantity changes within the same type from changes in types using wage data. In Section 6 below we explore this further in the context of a possible vintage "type" change with the introduction of computers.

The main identification approach used in this paper is the flat spot method which can be used with either homogeneous or heterogeneous models. The flat spot method, proposed in Heckman, Lochner and Taber (1998), is based on the fact that most optimal human capital investment models have the feature that at some point in the working life-cycle, optimal net investment is zero. The human capital of a given cohort over those years is constant. That is, there is a flat spot in the human capital life-cycle profile. Observing the changes in wages for the cohort over the flat spot, therefore, identifies the human capital price changes.

3 The Price Series for Human Capital

Many aggregate models of the economy implicitly use a homogeneous human capital model with an aggregate labor input and a single price (aggregate wage). However, the large microeconomic literature on increasing inequality in the U.S. and the links to increased estimated rates of return to schooling and skill biased technological change typically considers a framework with at least two skill levels. This is also an increasingly common feature in the macroeconomic literature which often divides the labor input into two types: skilled and unskilled.[8] In both literatures, the skill categories are defined with reference to observed education levels. For much of the literature high school graduates are compared with college graduates.[9] In this paper we estimate price series for four types of human capital or skills for the U.S. for the period 1963 to 2008. The four skill types, high school dropouts, high school graduates, some college and college graduates, are all defined with reference to observed education categories for ease of comparison with the previous literature.

3.1 Estimation Methodology

As noted above both homogeneous and heterogeneous human capital models assume an efficiency units structure at some disaggregated level. Under the assumption of competitive markets for each human capital type, log wages in these models for any individual i of a particular type are given by[10]

$$lnw_{it} = ln\lambda_t + lnE_{it}. \tag{2}$$

This implies that within each type the change in log wages between t and $t+1$ is given by

$$lnw_{it+1} - lnw_{i,t} = [ln\lambda_{t+1} - ln\lambda_t] + [lnE_{it+1} - lnE_{it}], \tag{3}$$

and, therefore, that the price change is given by

$$ln\lambda_{t+1} - ln\lambda_t = [lnw_{it+1} - lnw_{it}] - [lnE_{it+1} - lnE_{it}]. \tag{4}$$

Flat spot methods estimate the price change by restricting estimation to observations where human capital levels do not change over time, i.e. where $[lnE_{it+1} - lnE_{it}] = 0$, so that the price change is equivalent to the observable wage change. They assume that for each type there is an age range towards the end of the working life where efficiency units are constant. This is the point at

[8] See, for example, Krusell *et al.* (2000)

[9] In Krusell *et al.* (2000) skilled labor is defined "as requiring college completion or better (at least 16 years of school)" (p. 1033); the remainder are unskilled.

[10] Here we have suppressed the superscript notation delineating type.

which gross investment is just sufficient to compensate for depreciation, a typical feature of Ben-Porath based models of optimal human capital investment over the life-cycle.[11] Aggregating over these observations, the price series can be estimated from observed wage changes

$$ln\lambda_{t+1} = ln\lambda_t + [lnw_{t+1} - lnw_t] = ln\lambda_t + D_{t+1}, \tag{5}$$

where $D_{t+1} = [lnw_{t+1} - lnw_t]$ is the difference in log wages between t and $t+1$ for a sample of observations which have the same efficiency units in t and $t+1$.

Initializing $\lambda_0 = 1$, the price series is then estimated from successive log wage differences according to

$$\begin{aligned}
ln\lambda_1 &= D_1 \\
ln\lambda_2 &= ln\lambda_1 + D_2 = D_1 + D_2 \\
. &= \\
. &= \\
ln\lambda_T &= D_1 + D_2 + ... + D_T,
\end{aligned}$$

3.2 Data

The data source for the analysis is the annual March series from the Current Population Survey (MCPS). Repeated cross section data from the MCPS provide empirical counterparts to D_{t+1} as follows. Denote the age at the beginning of the flat spot as a. In year t the MCPS provides a representative sample of wage observations for individuals aged a in year t; it also provides a representative sample of wage observations for individuals aged $a+1$ in year $t+1$. Abstracting from mortality, these two observations provide a synthetic cohort. By assumption, efficiency units are constant for these individuals over the age range a to $a+1$, hence the difference in log wages in the sample of those aged $a+1$ in year $t+1$ compared to those aged a in year t, provides an estimate of D_{t+1}. Other estimates are provided by the difference in log wages in the sample of those aged $a+2$ in year $t+1$ compared to those aged $a+1$ in year t, or those aged $a+3$ in year $t+1$ compared to

[11] See Heckman, Lochner and Taber (1998), Huggett, Ventura and Yaron (2006) and Kuruscu (2006) for recent discussion of this feature of optimal life-cycle investment models.

those aged $a+2$ in year t, etc. Given a flat-spot interval of m years, there are $m-1$ comparisons that can be used in the estimation.

The MCPS records annual labor incomes for the year preceding the survey. Data from the March files for 1964 to 2009 were employed in the analysis to construct series covering earnings years 1963 to 2008. The MCPS has a number of advantages for this kind of analysis, most importantly the long time period and the representative nature of the sample. However, there are a number of important issues that arise in using these data. Two major concerns are the consistency in the definitions of the key variables, such as earnings, hours and education levels, over the period, and the time varying treatment of top-coded and allocated values. With regard to the former there is a break in the series in 1974-75 that affects the way annual hours can be constructed. Because of this break, price series are estimated separately for the 1963-1974 period and the 1975-2008 periods. There is also a change in the way education levels are recorded in 1991. This is dealt with using the evidence from a sample of workers covered by both definitions, detailed in Jaeger (1997) and discussed in the Appendix.

There are major changes in the treatment of top-coding and allocated values in 1988/89 and 1995/96 as well as frequent discrete changes in top-coding cutoffs over the 45 year period. For some broad aggregates, the changes have little effect, but for the price series estimated from experienced college graduates, for example, they can have very significant effects. As documented in the Appendix, problems with time varying treatment of top-coding and allocated values in the earnings are largely avoided if median wages are used in place of mean wages or the most commonly used log wages. These and other data issues, as well as the solutions adopted for this paper, and robustness checks are discussed in detail in the Appendix.

Our overall sample includes all paid workers between the ages of 19 and 64 who have positive earnings in the previous year. This large sample is used to construct aggregate labor quantities. Subsamples based on gender, education and age are used to construct the price series. Two alternative sets of these subsamples were used, based on total hours restrictions: the "unrestricted" sample requires only that individuals have worked at least five hours a week for at least 5 weeks last year; the "full-time-full-year" (FTFY) sample requires that individuals have worked at least 35 hours a week for at least 40 weeks last year. Hourly wages are constructed by dividing total annual earnings by total annual hours worked. The hourly wages are then deflated using the Consumer Price Index (1982-1984=100).[12]

[12]Most of the procedures dealt with in the paper are unaffected by the choice of deflator. The implied quantities of

3.3 Methodology for Choosing Flat Spot Regions

Standard human capital models predict three features of the life-cycle profile of human capital stocks that we use as guiding principles in generating the price series using the flat spot method. First, the profile is concave with the largest increase in the early years, rising to a peak and then declining as depreciation comes to dominate gross investment. This implies choosing a flat spot interval that is too early results in price change estimates that are upward biased. That is, they include some positive change in efficiency units, as net investment is still positive. Conversely, choosing an interval that is too late overestimates the price change because of the depreciation. Ideally the flat spot interval brackets the peak such that for the sample as a whole the change in efficiency units is zero. Second, the peak occurs at later ages for more educated groups. This suggests choosing a relatively late interval for the college graduate group. Third, post-school investments are generally predicted to be smaller, or even zero, for lower education groups compared to college graduates. Thus, the series for college graduates should be more sensitive to moving the flat spot interval than for high school dropouts.

The first step was to determine a flat spot range for the most sensitive college graduate group. The age ranges for the lower education (less sensitive) groups were then set by moving the age range back to keep the flat spot at the same experience point for each group. Earnings profiles from panel data cannot be used to identify the flat spot because of the basic price/quantity identification problem addressed in this paper. Similarly, in general cross section age-earnings profiles cannot be used to identify the flat spot due to cohort effects. The approach taken here is to use the cross section data for the college group for years when the bias from the cohort effects can be signed, based on the direction of selection inferred from Figure 1. There are two cohort effects that may bias the cross section profiles: different selection from the initial endowment distribution and technological improvement in human capital production functions. The cohort effects due to technological improvement impart a general downward bias on the slope of the human capital profile, since older ages in the cross section are from less advanced technology. The cohort effects due to selection impart a downward (upward) bias on the cross section slope in periods when older ages in the cross section are from cohorts with a higher (lower) fraction of college graduates (and hence of

human capital are derived by dividing the real wage by the estimated price series so the deflator cancels. Similarly, in the estimates of relative prices, the deflator cancels. However, the actual magnitude of (correlated) price movement over time is affected by the choice of deflator.

lower (higher) average quality).

The strongest downward bias over the potential age range for the earnings peak occurs for the 2003 cross-section. The relevant birth cohorts change from 1958 to 1946 as the cross section is aged over the age range 45 to 57, and, given the large increase in college attainment over these cohorts (see Figure 1), a negative selection effect is combined with any negative technological change effect. Moving forward, 2004 and 2005 should be similar to 2003, as the difference in college attainment from the relevant earlier and later cohorts stays the same. However, moving back towards 2000, 1999 and earlier cross sections should substantially remove the negative selection over the age range 45-57 since the relevant birth cohorts now change from 1955 to 1943, which have more equal college attainment (Figure 1), leaving only the negative technological change effect. Figures 2a plots the cross section age-earnings profile of FTFY workers for 1999 using the log of the median wage (Median), the log of the average wage (Wage) and the average of the log wage (Log Wage).[13] This cross section profile slope is assumed to be only minimally affected by negative selection bias around age 50. All three measures show the same pattern. The earliest age for any of the profile peaks is 55.

A more detailed investigation of the pattern of profile slopes over the range of cross sections from 1996 to 2007 reveals a pattern that is consistent with the expected selection effects from Figure 1. Table 1 reports some summary results in the form of a linear approximation to the slopes over ages 45 to 58 for the cross sections between 1996 and 2007. The slopes are all positive before 2001. As expected, the slopes flatten approaching the 2003-2005 cross sections, which are assumed to be subject to the strongest negative selection bias. The bias in the slope of the cross section profiles implied by Figure 1 vary over the age range. For the potential flat spot mid point range of 45-58, the 2003 cross section is subject to the largest bias. Detailed examination of Figure 1 indicates the downward bias should occur over the ages 45 to 53, since this is a move over the birth cohorts from 1958 (for age 45) to 1950 (for age 53). The birth cohorts from 1950 to 1946 have essentially the same fraction of college graduates so the bias from selection is removed at this point. Figure 2b shows the pattern of the 2003 cross section. Comparing Figure 2b with the 1999 cross section (Figure 2a), the primary difference is that while Figure 2a shows a generally positive slope throughout, Figure 2b shows an initial decline that is especially marked from 45 to 50. This is precisely the expected bias pattern by detailed age implied by Figure 1.

The evidence from the cross section age profiles suggests that the benchmark flat spot mid-point

[13]The profiles are similar for the larger unrestricted sample, and for adjacent years, 1998 and 2000.

for college graduates should be chosen no earlier than the mid-fifties.[14] In addition, the magnitude of the slopes shows that there is potential for significant downward bias over periods of a decade or two from using mid-points for the flat spot range that are too early. A typical slope in the period of cross sections 1996 to 2000 is around 0.8%; given the cumulative nature of the bias, this would produce a substantial bias after a decade.

A final requirement in order to implement the method is to choose the length of the flat spot, and hence the number of cohorts that can be used to identify price changes between any pair of years. There is a tradeoff between the length of the flat spot and the sample size. The analysis reported here uses a flat spot length of 10 years, allowing the averaging of 9 cohort pairs across any two years. This is, perhaps, the minimum length that is feasible given the sample sizes.[15]

3.4 Estimated Benchmark Price Series

Figure 3 plots the benchmark flat spot series. Because there is a break in the series between 1974 and 1975, corresponding to the change in the way hours data are recorded in the MCPS, Figure 3 presents both subseries with 1974 and 1975 each normalized to 1. The series are all based on wage observations for males.[16] The flat spot range of 50-59 for college graduates was chosen, based on the analysis of the cross sections above. The ranges for the remaining groups were chosen to keep the flat spot at approximately the same number of years of experience for each group: some college is 48-57, high school graduates are 46-55, and high school dropouts are 44-53.[17] The estimates in Figure 3 are derived from the FTFY sample and use median wages.

The most striking feature of Figure 3 is the close correspondence in the series for such diverse education groups as high school dropouts and college graduates. Over short intervals, the point

[14] The profile slopes for cross sections around 1999 are still subject to potential downward bias from technological improvement, even though the effect from selection is largely removed.

[15] Since the adjacent age pairs are averaged across all nine pairs corresponding to the flat spot age range, the primary requirement is that averaged over these pairs, the change in efficiency units is zero. It is not necessary that the change in efficiency units is exactly zero for all pairs in the range; indeed, it may well be the case that for the earliest ages in the range the efficiency units may still have a small increase, and for the oldest ages efficiency units may have begun to decline. Given the estimates of the human capital stocks implied by the series obtained from the assumed flat spot region, separate estimates of the stocks for each of the individual year age groups in the flat spot region can be computed. Simple formal tests cannot reject the hypothesis that these are equal.

[16] Females are excluded due to their larger fluctuations in labor force participation and the resulting added difficulty of defining an appropriate age range for their flat spot, as well as problems from time varying discrimination.

[17] This is a departure from Heckman, Lochner and Taber (1998) who used the same age range for their two groups - high school graduates and college graduates. Their assumption implies a later peak in net human capital investment in the experience profile for high school graduates than for college graduates, which is not a typical feature of standard optimal human capital investment models. However, we follow Heckman, Lochner and Taber (1998) in abstracting from the complication that the flat spot range may change over time.

estimates show some movement in relative prices by skill group, but any gaps tend to disappear quite quickly. The three highest education groups are closely related throughout. All four groups are closely related most of the time except for the pattern of the dropouts tending to suffer a larger price decline in recessions, though in all cases this is recovered within a few years. Relative prices are almost the same at the beginning and the end of the 45 year period. This is a surprising result in light of the large literature documenting and analyzing the increase in the rate of return to schooling, the relative wage of skilled workers, or the college premium, generally interpreted as an increase in the relative skill price.[18] The implication of substantial long term stability of relative skill prices together with changing relative wages is that the relative efficiency units of the different education groups has changed over time. In Section 6 below we argue that the sources of these changes are vintage effects caused by technological improvements in human capital production functions, broadly interpreted, and selection effects due to large changes in the distribution of education levels across cohorts.

The other main feature of Figure 3 is a pattern of substantial price change over time. From 1963 to the mid-seventies there is a price increase of 10 to 15 percent. From a peak in the mid-seventies there has been a major decline in the price, of over 15 percent. The pattern of the drop is interesting. It begins with a substantial decline until the recovery from the recession in the early 1980s. This is followed by a further decline until another period of recovery from the recession in the early 1990s. The price series for all of the education groups show the same broad sequence over all of these movements.[19]

3.5 Estimating Prices from Wage Data: Robustness and Sensitivity

In general, the use of wage data to estimate prices raises a number of problems, whatever the identification procedure. The main assumption employed in using wage data to estimate the prices is that workers are paid their marginal product in each period. Biases for some groups could arise from contract wages and time varying incentive effects. The standard literature associating relative wages with relative skill prices has discussed a number of these problems, including the complications

[18]Estimated relative prices are not directly reported in Heckman, Lochner and Taber (1998), but a pattern of an increasing skill price following skill biased technical change used in their simulation shows a larger relative increase in the price for college graduates compared to high school graduates than is apparent in Figure 3. As discussed in more detail below, this is largely due to their use of an earlier flat spot in the experience profile for college graduates than for high school graduates, which our cross section analysis indicates will cause an upward bias in the relative price.

[19]Heckman, Lochner and Taber (1998) estimated relative skill prices from de-trended data and therefore did not estimate a secular pattern of price increase or decrease for any group.

arising from changing non-wage benefits, union rents and other influences on relative wages. All identification approaches based on observed wages are subject to these problems.[20]

The benchmark series uses median wages. The use of medians has both theoretical and practical data advantages for a benchmark series. In terms of data issues, medians have the advantage of being much less sensitive to time varying treatment of top-coding, allocated values and outliers. A detailed analysis, given in the Appendix, shows that the basic pattern of the price series is insensitive to the use of alternatives to the median for all groups, provided problems arising from time varying top-coding methods are avoided. In practice, this is only a problem for the flat spot sample for college graduates.

In addition to the practical data advantages, the use of the median wage avoids a number of problems arising from departures from a consistent relation between wages and marginal products over time. For example, consider a population of homogeneous individuals in terms of human capital but variation in the form of contracts they are working under subject to the constraint that the discounted present value of the payment streams are equal. With all contracts having the form of spot wage equal to marginal product, the use of medians, log wages or average wages would yield the same price series. Suppose alternatively that a majority have a contract where the wage equals the marginal product every period, but that a substantial, and time varying fraction have contracts that deviate from this. The use of medians will continue to estimate the price series for this homogeneous group correctly, while the alternatives will estimate series that deviate from the true series. In periods of increasing shifts to alternative contracts, the deviation from the true series could be marked and long lasting. A similar argument can be made for changing union rents in situations where the median worker in terms of wages is not a union member, or for changes in some benefits which do not affect the median worker. The benchmark estimates use medians, based on these theoretical and data arguments, however, as described in detail in the Appendix, the basic pattern of Figure 3 is robust to alternative wage measures, sample restrictions and estimation methods.

A remaining issue is the sensitivity of the results to the choice of flat spot ranges. An independent check on the range for the lowest education group, high school dropouts, can be made by using a "standard unit" method that assumes that the human capital of young dropouts (i.e. with no experience) from the same point in the initial endowment distribution is constant over time, provided

[20]Possible non-random participation due to retirement is perhaps a more specific concern for flat spot estimates. However, the use of flat spot regions not too close to retirement, and the supplementary confirmation from an alternative method (discussed in the Appendix) that uses a younger age group, goes some way to minimize this problem.

the initial endowment distribution is stable. The full analysis is given in the Appendix. In summary, apart from a deviation following the recession of the early 1980s, not recovered until almost the end of the decade, the standard unit series is almost identical to the dropout flat spot series. (See Figure A1.) This close similarity obtained from the two independent methods - one using across cohort variation at a young age, and the other using within cohort variation at much later ages - provides some evidence of robustness of the flat spot estimates.

A similar comparison cannot be made for groups with higher levels of education, as their periods of schooling and significant post-schooling investment are longer making them unsuitable candidates for the standard unit method. Instead for them the pattern of sensitivity of the flat spot estimates to deviations from the benchmark age range was examined for consistency with the pattern implied by the cross section analysis of Section 3.3 and the cohort education paths in Figure 1. Figures 4-6 show the sensitivity of the flat spot series to changes in the flat spot age region for high school dropouts, high school graduates and college graduates.[21] The benchmark region for high school dropouts was 44-53. In Figure 4, dropout price series were plotted around this benchmark as 44/45-53/54 and 45/46-54/55 above the benchmark and 43/44-52/53, 42/43-51/52, and 40/41-49/50 below.[22] The results show that for dropouts the series are insensitive to the flat spot region. Moving the flat spot back to the earliest point shows a detectable tilting in the expected direction, but it is small suggesting that the true life cycle investment profile for dropouts is quite flat. Figure 5 repeats the analysis for high school graduates. As expected, this shows more sensitivity. There is an obvious fanning out, though the effect of moving to an earlier flat spot range is still relatively modest.[23] This is consistent with a steeper life cycle investment profile for high school graduates compared to dropouts.

Finally, Figure 6 reports the results for college graduates. Again the series fan out in the expected direction: moving the flat spot region to an earlier age tilts the profile upward. The sensitivity to moving the flat spot earlier is greater than for high school graduates. The analysis of Section 3.3 suggested that moving the midpoint of the flat spot for college graduates earlier than age 55 could cause an upward bias in the price series of around 0.8 of a percentage point each year. The magnitude

[21] Some college is very similar to high school graduates.

[22] Flat spot series were pooled across two adjacent series, e.g. 43-52 and 44-53, allowing for shifts of half a year in the mid-point of the region, and putting smaller weight on the ages near the beginning and the end of the range.

[23] For high school graduates there is more sensitivity in moving to a later age range; sensitivity in that direction may be due to the influence of pre-retirement behavior or depreciation.

of fanning out in Figure 6 is consistent with this. As noted earlier, this is why our estimates show a smaller increase in the relative price of college to high school graduates compared to Heckman, Lochner and Taber (1998). Figure 6 indicates that their use of a flat spot in the experience profile for college graduates that is earlier than for high school graduates, will cause an upward bias in the relative price.

4 Human Capital Prices, Inequality and the College Premium

The price series in Figure 3 show some differences across education groups. However, the most noticeable feature of the series is their high correlation. For example, the correlation between the high school graduate and college graduate series in Figure 3 is 0.91. This implies only modest changes in relative skill prices. How can this be reconciled with the well documented increase in the college premium in the 1980s and 1990s? The basic fact is that various summary measures of the difference in wages or annual earnings for those with a college degree compared to, say, high school graduates did increase substantially. For example, Card and Lemieux (2001, p.705) report an increase in the gap "from about 25 percent in the mid-1970s to 40 percent in 1998." The standard approach to analyzing this increase in inequality is to posit a heterogeneous human capital model in which college graduates are one type of human capital and high school graduates are another type and to attribute the change in the gap entirely to a change in the relative prices of these two types of human capital. However, as discussed in Section 2, this implicitly imposes the strong identification assumption of no change in the relative quantities. It rules out both technological change in human capital production functions and selection effects which would be expected over periods of substantial changes in cohort education levels. Moreover, it implies that the path of the college wage gap should be the same for all ages since all ages would be subject to the same relative price changes, which is strongly inconsistent with the data.

There is, in fact, a strong age pattern to the recent increase in the college wage gap documented in Card and Lemieux (2001), hereafter CL, using census data. Figure 7 shows the evolution of the college premium for males from the MCPS data used in this paper for the two age groups, 26-30 and 46-60, that were used in Figure 1 in CL. It plots the log of the ratio of median hourly earnings of college graduates to high school graduates and shows a very similar pattern to Figure 1 in CL.[24] The

[24]The sample is full time and full year workers, but the same pattern holds for the less restrictive sample requiring at least 5 weeks of work with at least 5 hours a week.

recent premium increase is more important for younger workers (26-30), where the ratio declines slightly before 1980, then increases. It increases more slowly for the older workers, particularly in the 1980s.

Using our estimates, the log of the ratio of median hourly earnings of college graduates to high school graduates, plotted in Figure 7, can be decomposed into relative quantity and price components as follows

$$ln(\frac{w^c}{w^h})_t = ln(\frac{\lambda^c}{\lambda^h})_t + ln(\frac{E^c}{E^h})_t,$$

where $\frac{w^c}{w^h}$ is the ratio of wages of college graduates to high school graduates, $\frac{\lambda^c}{\lambda^h}$ is the ratio of prices and $\frac{E^c}{E^h}$ is the ratio of quantities. Figure 8 presents this decomposition for the young age group (26-30). The movement in the wage ratio is closely tracked by the movement in the efficiency ratio. The relative price path in Figure 8 uses the price series from Figure 3. It shows a flat or very slightly decreasing relative skill price to the late 1970s, an increase to the mid or late 1990s, followed by a decrease to 2008. Throughout, the movement in relative prices is relatively modest, reflecting the high correlation of the price series for the different education groups Figure 3. The decomposition shows that the relative quantity variation is generally more important than the relative price variation. Most of the rapid rise in the relative wage for young college graduates over the 1975 to 1995 period comes from the increase in the relative quantities. From 1975 to 1995 there is a change in the relative log wage ratio of .262. The decomposition shows that .178 comes from the quantity change and .085 from the price change. Thus, the price change only accounts for one third of the observed wage premium change: two thirds is due to the relative quantity change.

The birth cohorts for the older age group in CL ranged from 1910-1924 for the 1970 observation to 1935-1949 for the 1995 observation. Thus, Figure 1 indicates that for most of this period there is a negative ability selection effect implied by the increasing fraction of college graduates in successive cohorts. This only reverses for the older group after the early 1990s. Thus, by contrast with the younger group, there is no strong positive selection effect in the 1980s, implying faster growth for the younger group in that period. Figure 7 (and CL Figure 1) in fact shows that the faster increase in the college premium for the younger workers does occur mainly in the 1980s. In the 1990s, the rates of increase are more similar. This is also consistent with the expected selection effects implied by Figure 1 since by the 1990s the differential selection effects across the young and old groups that occurred for the 1980s are largely removed.

Overall this decomposition indicates that relative quantity changes are more important than relative price changes in explaining the observed changes in relative wages. While there is an important role for relative prices changes, the common assumption that all wage differentials are driven by price differentials is not supported by the evidence and results in misleading conclusions. The next section examines the pattern of life-cycle human capital profiles by education group across successive birth cohorts. It reveals more detail on the implied magnitudes of the changes in the quantities of human capital within education groups across cohorts due to the selection and technological improvement effects that are behind the relative quantity changes during the period of the increasing college premium.

5 Human Capital Prices and Life-Cycle Analysis

The fundamental identification problem in human capital models discussed in Section 2 presents a major problem for interpreting life-cycle age-earnings profiles. In the standard life-cycle human capital model of the Ben-Porath type, observed wages are the product of a price and (supplied) quantity of human capital. Identifying the life-cycle profile of the (supplied) quantity of human capital from wage data requires identification of the price. Even with cohort data, aging a cohort over time does not identify the time profile of a worker's supplied human capital unless the price is constant over the lifetime. In almost all of the literature on life-cycle earnings a constant price is a maintained assumption.[25] For example, the relevant chapter in the *Handbook of Labor Economics* has no discussion of time varying prices.[26] Under the constant price assumption the life-cycle wage profile is the same as the life-cycle (supplied) human capital profile.[27] The pattern of life-cycle wages can then be used to directly test human capital model predictions concerning the life-cycle profile of human capital.

Kuruscu (2006) is a recent example of life-cycle analysis that takes this approach. Marginal

[25] The main exception is Heckman, Lochner and Taber (1998). More recently Huggett, Ventura and Yaron (2006) relaxed the constant price assumption and assumed a constant rate of growth for the rental rate on human capital equal to the average growth rate in mean real earnings. Guvenen and Kuruscu (2007) take a similar approach in calibrating their model of inequality.

[26] See Weiss (1986).

[27] In the Ben-Porath model, in each period of positive investment the individual supplies part of their human capital to the market and part to the production of more human capital. After the period of specialization, the part supplied to the market may be the largest part, and in some cases, almost all the human capital stock of the individual at that point, but in general the supplied and total stocks differ until gross investment is zero.

costs of post-school investment, based on the Ben-Porath model, are estimated with the explicit assumption of a constant rental rate over the life-cycle. This assumption is used to infer from wage profiles that growth in human capital over the life-cycle stops relatively early and leads to the conclusion that training has a small contribution to lifetime income. However, the estimates presented in Section 3 strongly indicate that the rental price is not constant, and that a constant price assumption will lead to misleading conclusions about the life-cycle profile of human capital. In fact, the evidence suggests that in the last three decades in the U.S. the price movements have been large.[28] In this section we first compare the implied life-cycle human capital profiles for a variety of birth cohorts whose wages are observed in the 1963-2008 period using the price series of Section 3 with the implied profiles using the standard constant price assumption in the literature. We then examine the pattern and magnitude of the estimated human capital quantity changes across cohorts and relate these to the vintage effects implied by selection and secular technological improvement effects.

Under the constant price assumption the life-cycle (supplied) human capital profile is the same as the life-cycle wage profile. This is plotted for males in the lowest skill level, high school dropouts, for selected cohorts spanning the earnings observations in the MCPS for 1963 to 2008 in Figure 9a.[29] The profiles are difficult to make sense of within a standard Ben-Porath model. They are all very different shapes. The 1925 cohort appears to have continued to grow quite rapidly to age 50; the 1937 cohort shows extremely rapid growth to a peak around age 40. In contrast, the 1946 birth cohort shows rapid growth in the twenties but peaks around age 30 and the 1958 cohort is flat throughout. The profiles often cross. Moreover, the most recent cohorts show low levels of human capital relative to the earlier cohorts. Some decline could be expected between the 1925 and 1937 cohorts due to selection. In Figure 1, the fraction of the high school dropouts in a cohort shows a substantial decline up to the 1946 cohort. Given a positive correlation between initial endowment/ability and completed education,[30] the decline in the cohort fraction of high school dropouts would be accompanied by a decline in the median initial endowment/ability among the high school dropouts up to the 1946 cohort. However, after 1946 the fraction is stable. It is, therefore, surprising that the 1958 cohort

[28] The evidence also suggests that the movements in the price have not been monotonic. Thus assuming a constant growth rate in the price, as in Guvenen and Kuruscu (2007), would also result in misleading conclusions.

[29] All of the life-cycle analysis is done for males only. A full analysis for females needs to deal with the selection arising from a different life-cycle participation pattern.

[30] See footnote 7.

appears to have so much less human capital than the 1946 cohort when it should be drawing from the same point in the initial endowment distribution.

Figure 9b shows the implied life-cycle human capital profile for the same group using the price series for high school dropouts from Section 3 to identify the profile. Even though the profiles are plotted with no smoothing in any of the underlying series, it is apparent that the pattern is now much closer to a series of cohort profiles all with similar shapes and more readily interpretable within a Ben-Porath model with slow changes in production function parameters and/or initial endowments. Instead of drastically varying shapes in the early twenties to early thirties age range, and drastically varying peaks from age 30 to age 50, the profile shapes are much closer to each other and to a standard concave profile. There remains some indication of a small drift down over time in the profiles. The 1958 cohort still appears somewhat below the 1946 cohort, but compared to Figure 9a the gap is much smaller and the slopes are quite similar.

Figures 10a and 10b repeat the analysis for high school graduates; while figures 11a and 11b plot the estimated profiles for some college. The same contrast appears as for high school dropouts, though the pictures are even clearer due to the larger sample sizes which make the profiles smoother. Figures 10a and 11a shows the same confused pattern as Figure 9a. There is a lot of crossing in the profiles and the 1958 profile is dramatically worse than the 1946 profile, with a twenty to thirty percent difference for most ages up to the early 40s for high school graduates. In contrast, Figures 10b and 11b show the classic Ben-Porath profile shape for all birth cohorts and much more similarity in the 1946 and 1958 cohorts.

Finally, Figures 12a and 12b repeat the analysis for college graduates. In this case, the benchmark series again produces more similar shapes and eliminates significant crossing in the profiles, though the contrast with the standard constant price assumption is less apparent. An important difference is that the most recent 1958 profile shows a consistent improvement over 1946 using the benchmark series which does not occur with the constant price. This is pursued in more detail below.

6 Vintage Effects: Interpretation and Discussion

Overall, the use of the price series from Section 3 provides a picture of cohort change over time in the human capital profiles that is much less erratic, and much easier to explain in an optimal human capital life-cycle investment model with moderate changes in the production function parameters over time. The decomposition in Section 4 showed an important role for changes in the quantity of human capital for college graduates of different vintages. The different vintages of college graduate may have different "types" of human capital or different amounts of the same "type". The maintained hypothesis explored in this paper is that college graduates have the same "type" of college graduate human capital over this period and therefore that profile shifts reflect different amounts of the same type of human capital associated with different vintages of college graduate. In this section we contrast this hypothesis with an alternative skill biased technological change hypothesis under which there was a change in vintage "type", coinciding with the widespread introduction of micro computers, that resulted in an increase in the college premium for the young ("new vintage") college graduates, but not for the ("old vintage") older college graduates, as described in Card and Lemieux (2001).

The primary difference between the two hypotheses is their different implications for the time path of vintage effects. Under the hypothesis explored in this paper vintage effects occur through selection effects and technological improvement in human capital production which occur continuously. Within an age/sex cell for college graduates the quantities will change continuously through these vintage effects. The alternative hypothesis is that vintage effects come about through changes in vintage types rather than through quantity changes due to selection effects and technological improvement. Within an age/sex cell, there is no change in quantity for the same vintage type.[31] Thus, for the maintained hypothesis of this paper, the time path of the vintage effects is constrained by the selection effects implied by the cohort education patterns in Figure 1 combined with secular trend improvement in the production of human capital. For the alternative hypothesis the pattern is constrained by the timing of the vintage type change.

The general pattern of vintage effects for college graduates that follows from the maintained hypothesis is provided in graphical form in the life cycle profiles found in Figure 12b. Figure 1 shows a substantial increase in the fraction of college graduates from the 1931 to the 1946 birth cohorts, the positive correlation between ability and education level suggests a negative ability selection over this period, followed by a reversal of the selection effect as the fraction began its decline until the most recent cohorts. After the 1946 birth cohort, both selection effects and technological improvement predict an upward shift in the profiles. To get a more detailed picture, and an indication of the magnitude of the shifts across cohorts, the cohort effects in Figure 12b were approximated by

[31] In fact, if there is a change in vintage type then the quantities are not directly comparable.

imposing a quadratic specification in age (experience) over the same mid career age range (30-45) for each the relevant (three year) birth cohorts and estimating the cohort intercepts. The results are shown in Table 2. The omitted cohort in Table 2 is the 1946 cohort (1945-1947) which, in terms of selection effects, is the turning point.

The pattern of cohort intercepts, relative to the 1946 birth cohort is exactly what would be expected from constant slow technological improvement together with the selection effects implied by Figure 1. Moving forwards from the 1946 birth cohort by assumption the technology effects are positive for all the subsequent cohorts, and from Figure 1 for the 1949 to the 1958 cohort, the selection effects are positive, while from 1946 to 1949, and from 1958 to 1961 they are zero. The cohort dummy variables reflect this pattern with relatively large increases between 1949 to 1958. The combined magnitude from the "worst" post war cohort of 1946 to the "best" cohort of 1961 is 10.8 percent over 15 years. In 1946 30.47 percent of the cohort became college graduates; by 1961 this had fallen to 24.62 percent, implying a potentially large selection effect as up to 20 percent of the worst students are no longer going to college.

Moving backwards from 1946, the selection effects imply improved cohorts, but unlike the move forward from 1946, this will not be augmented by secular technological improvement, but rather may be partially offset by the negative effect of moving to earlier human capital production functions. Thus, in moving to the cohorts before 1946 the improvements should be reduced relative to moving forward. The estimates show this: the earlier cohorts moving backwards from 1946 are better, but by more modest amounts than the later cohorts moving forward from 1946. The 1943 and 1958 cohorts have roughly similar fractions going to college and therefore differ mainly because of technological improvement. The magnitudes suggest secular technological improvement at an annual rate of about a third to half a percentage point. This suggests that the 10.8 percentage point improvement in the 1958 cohort over the 1946 cohort is composed of roughly equal contribution of selection effects and technological improvement.[32]

The potential magnitude of the selection effects can be assessed from an examination of the distribution of human capital quantities within the 1946 birth cohort which had the largest fraction of the cohort completing college. The average real hourly wage in the FTFY sample for males between 40 and 49 years of age, belonging to the 1946 birth cohort is $16.06. If there was a stable correlation

[32] As an alternative, the cumulative difference in the profiles (compared to the 1946 profile) over 30-45 was estimated without the quadratic approximation and produced the same pattern.

of one between ability and education, an approximation to the magnitude of the selection effect for a given cohort can be made by comparing this unconditional mean with the conditional mean after selecting out the appropriate fraction of the population from the lower tail of the distribution. Selecting out 10% results in a conditional mean that is 7.34% higher; 20% results in a mean 13.93% higher, and 25% results in a mean 17.26% higher. The (negative) difference in the fraction of college graduates in the 1937 and 1946 birth cohorts is about 37%, and between the 1958 and 1946 birth cohorts is about 21%. The predicted human capital difference due to selection for the 1958 cohort over the 1946 cohort is 14.60%, and for the 1937 cohort is 25.81%. Given the assumption on the correlation these are upper bounds. In Table 2 the human capital of the 1958 cohort is 8.65% above the 1946 birth cohort, which includes both selection and technological improvement effects. Attributing one half of this to selection is quite consistent with a positive correlation between ability and education of much less than one.

Overall, the predicted timing and pattern of the vintage effects under the hypothesis explored in this paper is confirmed by the data. The predicted timing and pattern of the vintage effects for the alternative hypothesis depends on the assumptions made regarding the timing of the initial change in the vintage type, and the pattern of its spread to all new college graduates. Under this hypothesis there was a change in vintage "type", coinciding with the widespread introduction of micro computers. The timing of this change is typically taken to be in the mid 1970s.[33] Assuming a start date of 1975, if the new vintage type was produced at college, then the first new vintage cohort on the market - those trained to work with the new technology - is the 1957 birth cohort. If the new vintage type is some combination of college production and labor market entry of college graduates that can be trained in the new technology at work, then the first new vintage cohort is the 1955 birth cohort. In terms of the three year cohorts in Table 2, cohorts up to 1952 are the "old vintage" and the price series in Figure 3 now only applies to these cohorts. The new vintage has a higher price than the old vintage. Thus, applying the price series in Figure 3 to college graduates after the 1952 (or 1955) birth cohort will overestimate the quantity of the new vintage.[34]

The two hypotheses have similar predictions for the post 1952 birth cohort intercepts in Table 2, but have contrasting predictions for the earlier cohorts. The hypothesis explored in this paper

[33]This is the timing used, for example, in Heckman, Lochner and Taber (1998).

[34]A flat spot price series for the new vintage type cannot be made until the new vintage cohorts are in the flat spot age range. Using a mid-point of 55, the mid-point will not be in range until 2010, so the price series for this type cannot be estimated until after 2010.

predicts a "U" shaped pattern for the intercepts between the 1937 and 1952 birth cohorts with the relative magnitudes as described above. The alternative hypothesis predicts a flat profile since all these cohorts are the same vintage type. The results in Table 2 are inconsistent with the alternative hypothesis.[35] Both hypotheses predict that the profiles after 1952 should be higher. The hypothesis in this paper predicts both higher profiles in general and that each successive cohort up to at least the 1961 cohort will have a higher intercept in Table 2, but the rate of increase will slow as the selection effect changes from positive, to neutral and eventually to negative. This is the pattern that appears in Table 2. The alternative hypothesis predicts that the intercepts for the cohorts after 1952 will be higher in general due to the fact that the price series estimated on the old vintage is being used, and this will over-estimate the quantity for these cohorts. Without further assumptions, the alternative hypothesis has no predictions for the specific pattern of intercepts after 1952. Predictions for this pattern will depend on time path of the fraction of the post 1952 cohorts that are actually new vintage, i.e. on the pattern of diffusion of the new vintage.

7 Human Capital and Growth: Measuring the True Labor Input

Assessing the contribution of human capital to growth requires a measure of the labor input or inputs. The earliest measures of the labor input used in the literature were aggregate hours of labor. However, hours of different workers are typically not comparable in terms of their labor input. A worker with more human capital supplies more labor input. Several countries have attempted to get better measures by refining the construction of their aggregate labor input indices, previously measured by aggregate labor hours, to take into account changes in the composition of the labor force. The rapidly increasing average education levels in the workforce was a major reason for this initiative.

A homogeneous human capital model has great benefits as the conceptual basis for defining an aggregate labor input and corresponding aggregate wage. The single price feature of the model provides an elegant solution to the definition of the aggregate wage: the price of an efficiency unit of homogeneous human capital. The single type feature provides a similarly elegant solution to defining the aggregate labor input: the quantity of efficiency units of human capital. This quantity can be obtained by a simple aggregation of hours supplied weighted by the efficiency units of each worker.

[35]An F test of equality of the intercepts in Table 2 strongly rejects the null.

The high correlation exhibited by the price series in Figure 3 implies that for aggregate level analysis, the homogeneous price assumption can be used as a reasonable approximation. In this section, we examine the construction of measures of the true aggregate labor input using this approximation, and compare the results with the standard aggregate measures in the literature.

The issue is quite different if human capital is heterogeneous. If human capital is heterogeneous and the types are observationally identifiable by, say, education level, there is in fact little to be gained by arbitrarily aggregating the different types. One could just directly model the production function with all of the different human capital types.[36] In this case, calculating the true labor input of each type can be accomplished by using the separate price series in Figure 3 in place of a single price.

7.1 Aggregate Labor Input: Composition Adjusted Hours

Most standard aggregate labor input measures are some form of composition adjusted hours. The BLS provides the main official composition adjusted series for the U.S. as part of its Multi-factor Productivity (MFP) Program. The motivation for the series is described in BLS Bulletin 2426 (1993). Prior to this series, labor input had been measured by the total hours of all workers. It was widely recognized that "the effective quantity of labor input does not rest solely on the total number of hours worked by members of the U.S. labor force but also on characteristics of the labor force." (p. iii). Following the recommendations of a National Academy of Sciences Panel to Review Productivity Statistics in 1979 the BLS developed a weighted measure of total hours focusing on the skill level of workers as reflected in education and job market experience levels. This measure is used in the construction of the BLS MFP index.

The BLS measure is described in detail in the BLS Handbook of Methods (1997), and in BLS Bulletin 2426 (1993), which reported the first estimates. It is based on a Tornqvist chained index of weighted hours of workers classified by skill and demographic characteristics. The hours measures used in the original BLS Bulletin 2426 (1993) study for the period 1968-1990 were obtained from the MCPS. For the current series for the BLS MFP Program, hours are obtained mainly from the BLS Current Employment Statistics (CES) program, based on establishment surveys. They are supplemented by data from the CPS and other sources for groups not covered under CES.

[36]A recent example of this is Johnson and Keane (2008) who estimate a model with 160 different types of human capital differentiated by gender, education, age, and occupation.

The weights are the shares of total compensation for each type of worker classified by skill and demographic characteristics, and the weights are allowed to vary each year.

Prior to the development of the BLS measure, a number of authors had developed and published composition adjusted aggregate hours series.[37] The most well known current version of these is the Jorgenson series for the U.S. private economy, 1977-2000.[38] There are some differences in the details of the methods and coverage, but the basic methodological approach is the same for both the Jorgenson and BLS series, and the two series are very similar for the 1977-2000 period.[39]

The growth in the labor input series that adjusts for composition is substantially higher than the aggregate hours growth. Using the BLS figures in Bowlus and Robinson (2008) Table 2, the hours growth is 50.42 percent, but the composition adjusted input growth is 66.68 percent. Thus, the changing composition contributed almost one quarter of the total growth in the composition adjusted input. Since the growth rate using composition adjusted hours is almost one third higher than when using hours, the use of hours in constructing the MFP would substantially bias the change in the index over this period. Adjusting hours for composition changes is clearly important. However, because it ignores technological change in human capital production and endogenous choice of human capital investment, the composition adjusted series itself is subject to bias.

In addition to the Jorgenson series and official aggregate labor input measures, other measures of aggregate input, focusing on composition adjustment, have been constructed in a variety of studies in the business cycle literature and the macroeconomics literature more generally. Studies of wage cyclicality, recently reviewed in Bowlus, Liu and Robinson (2002), are concerned with the effects of a downward composition bias on the estimates of the correlation between wages and the labor input over the cycle. In tackling the problem of composition bias, these studies implicitly or explicitly construct aggregate wage and hours measures that are designed to address quality variation in the human capital input over the cycle induced by composition changes. Examples of include Hansen (1993) and Kydland and Prescott (1993) for a total economy aggregate, and Katz and Murphy (1992) and Krusell et.al. (2000) for aggregates by skill group. These series are all efficiency units based, either for the economy as a whole or within skill group. They all use a composition adjustment approach, and as shown below they produce very similar estimates to the BLS and Jorgenson

[37] See, for example, Chinloy (1980), Denison (1985), and Jorgenson, Gollop and Fraumeni (1987).

[38] Available at: http://post.economics.harvard.edu/faculty/jorgenson/papers/lqualprivate.xls

[39] The series are compared in detail in Table 2 in Bowlus and Robinson (2008). When the Jorgenson series is scaled to the BLS series in 1977, the two labor input series look almost the same.

estimates, to which they are closely related.

7.2 Aggregate Labor Input: Efficiency Units

With homogeneous human capital a total efficiency units series can be calculated simply by dividing total wage payments by the (single) price series. In this section an efficiency units series is compared with standard aggregate labor input measures. For ease of comparison, we use a single data set, the MCPS, to construct three measures of the labor input: an aggregate hours measure, a BLS/Jorgenson style composition adjusted aggregate hours measure, and the efficiency units measure implied by a single price series based on the estimates presented in Section 3.[40] Comparisons between these three measures have a simple interpretation within the homogeneous human capital model. Divide hours of labor into J types (skill groups) where within type all members are the same "quality," i.e. have the same efficiency units per hour. For the BLS these are groups based primarily on sex, education and experience.

Let E_{jt} be efficiency units per hour for a member of group j, and h_{jt} be the total number of hours of type j. Total efficiency units in period t, N_t, are then

$$N_t = \sum_j N_{jt} = \sum_j E_{jt} h_{jt}.$$

Let W_{jt} be total payments to members of group j in year t. Total efficiency units for any group j can then be computed simply by dividing the total expenditure, W_{jt}, by the price, λ_t, estimated in Section 3. Aggregating across types, the total efficiency units series N_t, is simply total expenditure across types, $W_t = \sum W_{jt}$, divided by λ_t.

The BLS measure uses a Tornqvist chained index. The construction of a Tornqvist chain index of composition adjusted labor input is as follows. For group j the ratio of the labor input in year t to the input in $t-1$ is by definition: $L_{j,t/t-1} = E_{jt}h_{jt}/E_{jt-1}h_{jt-1}$. Aggregating across groups, the Tornqvist chained index (ratio) of the total labor input in year t to the input in $t-1$ is given by weighting the ratios of the groups as follows

$$L_{t/t-1} = (L_t/L_{t-1}) = \prod_j (E_{jt}h_{jt}/E_{jt-1}h_{jt-1})^{\omega_{jt}}$$

[40]The full sample is used for these calculations including both males and females.

or

$$lnL_{t/t-1} = \sum_j \omega_{jt} ln(h_{jt}/h_{jt-1}) + \sum_j \omega_{jt} ln(E_{jt}/E_{jt-1}), \tag{6}$$

where the weights, ω_{jt} are the shares of the groups' efficiency units in total efficiency units, averaged over the adjacent periods

$$\omega_{jt} = (W_{jt}/(\sum W_{jt}) + W_{jt-1}/(\sum W_{jt-1}))/2.$$

The BLS series from an initial period zero to t follows by chaining the ratios, $L_{t/t-1}$, to get the change from zero to t

$$\Delta L_{t/0} = L_{t/t-1} L_{t-1/t-2} .... L_{1/0}$$

so that the value in any period t is given by

$$L_t = L_{t/t-1} L_{t-1/t-2} .... L_{1/0} L_0,$$

where L_0 is some normalized value in period zero.

Note that the BLS series implicitly assumes that there is no change in efficiency units per hour within groups, i.e. that $E_{jt} = E_{jt-1}$, which sets the second term in equation (6) to zero. The first term in equation (6), the BLS measure, is simply a composition adjusted hours index: a weighted sum of each groups hours change. Thus, the series ignores the second term which is the weighted sum of the percentage changes in average efficiency units per hour, or quality, within group. This term is non-zero whenever any group has a change in average efficiency units per hour over time, i.e. an average quality change via technological change, selection effects or cohort optimal investment changes.

While the BLS and Jorgenson methods use chained indexes of weighted hours growth rates with varying weights, the fixed weight methods with heterogeneous human capital, such as Krusell et.al.(2000), simply compute the hourly efficiency units of a worker of any given type as the average hourly wage of workers of that type. Applying the fixed weight efficiency unit method to all workers yields a total labor input, I_t, in period t of

$$I_t = \sum_j (\bar{W}_j h_{jt}), j = 1, 2, ...J,$$

where h_{jt} is total hours of workers of type j in year t, $\bar{W}_j$ is the average wage of workers of type j in the reference year (or averaged over all years), and J is the number of worker types. Since this is basically a composition adjustment approach with different weights, it suffers from the same type of bias as the BLS and Jorgenson estimates.[41]

7.3 Aggregate Labor Input: Comparison of the Alternative Measures

Table 3 compares the alternative labor input series estimated using the MCPS. The BLS-style (Tornqvist) composition adjusted series was calculated as described above, using 120 groups classified by education, age and sex for a population of paid workers aged 20-64.[42] The first column reports the aggregate hours estimate from the MCPS. The growth in hours is substantially less than the growth in the composition adjusted hours which are reported in the third and fourth columns. However, the growth in composition adjusted hours is itself substantially less than the growth in efficiency units reported in the final column of Table 3. Composition adjusted hours grow faster than the unadjusted series because of the increased education level in the population. Efficiency units grow faster than composition adjusted hours because the composition adjustment ignores technological change and changing investment patterns.

The magnitudes of the differences are large: efficiency units grow almost twice as fast as hours. The magnitudes of the differences in the growth rates are shown in Table 4. Unadjusted hours of paid workers for the period 1977 to 2000 has a growth rate of 60.94 percent; composition adjusted hours grew by 80 percent. The composition adjustment thus produces a labor input growth that is about one third higher than the unadjusted hours growth. However, the growth in efficiency units is 100 percent. The standard composition adjustment to hours is therefore only half of the full adjustment to aggregate hours that is necessary to estimate labor input growth between 1977 and 2000. The pattern is the same for private sector paid workers.

To examine the magnitude of the underestimate of the labor input using the fixed weight efficiency units methods we constructed efficiency unit aggregates by skill and in total using a method analogous to Krusell et. al. (2000) and Kydland and Prescott (1993). As noted above, these fixed weight methods are similar to the BLS and Jorgenson methods in that they aggregate the hours of different types of workers using average wages as weights, classifying the different types of workers according

[41] Krusell et. al. (2000) use the weights from the year 1980 for the whole period 1964-1993; Kydland and Prescott (1993) use the weights from averaging across all years; BLS and Jorgenson use time varying weights described above.

[42] The March supplement weights were used for all the total estimates.

to age, sex and education. The composition adjustment applied to aggregate hours implied by the fixed weight approach is almost identical to the estimates obtained for BLS style methods.[43] For paid workers, for example, Table 4 shows a growth of 79.86 percent from 1977 to 2000 using the BLS method; the fixed weight method estimate for the same period is 80.21 percent. Thus, the fixed weight method has the same degree of underestimation of the increase in the labor input as the BLS method. Fixed weight methods, by construction, do not permit total efficiency units of labor to increase if the demographic composition does not change, except through hours. This likely has little effect for cyclical analysis, but for longer term secular growth or cross country comparison, it is potentially extremely important. One important consequence is the potential for serious overestimation of MFP and underestimation of the role of human capital in growth.

Table 4 also reports the growth rates of alternative labor input measures by sex. The BLS method for total hours uses compensation shares to weight the growth of each type of hours, including male versus female. The logic of this weighting suggests that to get separate totals for males and females, the total labor input estimate should be split between males and females according to the compensation shares in the year, assuming no discrimination.[44] The results for this method are denoted BLS (A). An alternative is to apply the BLS method separately to estimate compensation share weighted male hours growth and compensation share weighted female hours growth. The results in this case are denoted BLS (B).[45]

Human capital theory predicts that the increased labor market attachment of females has increased female human capital investment. The substantial literature on female wage differentials has documented this increase, which has taken many forms, including more market oriented human capital investments for females at college. This increase has resulted in an increase in the total labor input of females by all measures, including total hours. Total hours for female paid workers increased by 90.98 percent from 1977 to 2000, which is double the growth in male hours of 42.75

[43] See Bowlus and Robinson (2008) for full details.

[44] The use of compensation shares in the BLS method implicitly assumes that the wage rate for females reflects the true marginal product, i.e. that there is no discrimination. The estimates of total efficiency units in Table 3 are also based on this assumption. If discrimination creates a significant difference between the wage and the marginal product of female labor, without adjustment the total efficiency units series would be underestimated, and the degree of underestimation would vary over time as the degree of discrimination varied. In a standard employer discrimination model the true efficiency series is calculated separately for males and females. For males it is calculated as before by dividing total wage payments by the estimated price; for females, the total wage payments first have to be scaled up according to the amount of the discrimination. If, for example, discrimination against females was declining over the period, the growth in efficiency units for females would be over-estimated.

[45] By construction, the relative rates of growth in the BLS (B) measures of the labor input by sex simply reflect the relative rates of growth of hours.

percent. The same pattern occurs for private sector paid workers. The growth in efficiency units (EUS) for females, however, is particularly pronounced. From 1977 to 2000 the growth in efficiency units for females is 172.17 percent, which is almost double the growth in hours. In contrast, much smaller rates of growth are estimated using the BLS style measures: 142.96 percent for BLS (A) and 122.01 percent for BLS (B).

7.4 Consequences for Multi-factor Productivity

A major motivation for the construction of quality adjusted labor input series is that the use of unadjusted hours results in a substantial bias in the estimation of MFP or TFP. Since changes in MFP are defined as the residual change in output that cannot be accounted for by the changes in the inputs, the estimates of these changes depend on the estimates of the changes in the inputs. Define l as the growth in the true labor input, h as the growth in aggregate hours, and h^c as the growth in composition adjusted hours. Then the overestimate of the growth in MFP from using h in place of l is $s_l[l - h]$ and the overestimate of the growth in MFP from using h^c in place of l is $s_l[l - h^c]$ where s_l is the share of labor in total costs.

The results in Tables 3 & 4 indicate that adjusting for composition falls short of a full quality adjustment, since it cannot capture technological change in human capital production or increased human capital investment by females. For the U.S. for the period 1975 to 2001, the growth in hours underestimates the growth in efficiency units of paid private workers by 62.04 percentage points. Since the share of labor in total costs is roughly two thirds,[46] this implies that MFP would be overestimated by about 40 percentage points if unadjusted hours were used. Using composition adjusted hours makes a substantial correction to this, but still underestimates the growth in efficiency units by 36.46 percentage points. Hence, this BLS type adjustment still implies an overestimate of the growth of MFP of 24 percentage points. The actual BLS estimate of MFP growth in the private business sector between 1975 and 2001 is 23.76 percent.[47] The results therefore suggest that all of this could be due to an undercount of the increase in the labor input. In fact, as noted earlier, our efficiency units estimate for females is probably too high which exaggerates the underestimate of efficiency units and the overestimate of MFP . For example, the overestimate of MFP would be reduced to below 20 percentage points if discrimination against females over the period declined in

[46]The BLS estimates for labor share in total cost are 0.678 in 1975 and 0.686 in 2001.

[47]See Table PB4a in mfp2ddod.txt at the BLS Multi-factor Productivity website.

the range of 10-12 percentage points.

Even with a reasonable adjustment for the estimated increase in female efficiency units, these results for MFP indicate that much of the source of improvement over time in standard of living is due to technological improvements in the production of human capital or increased human capital investment. Individuals exposed to more recent education and on-the-job training systems receive more value added to their human capital. This is not captured by composition adjustment. In particular, composition adjustment cannot capture a change in the level of human capital accumulated by college educated workers from the 1966 birth cohort compared to the level accumulated by an otherwise identical individual from the 1946 birth cohort. Similarly, composition adjustment cannot capture the increased human capital for females that would be expected from a large increase in lifetime participation and hours for females.

8 Discussion and Conclusion

Separate identification of the price and quantity of human capital has important implications for understanding key issues in labor economics and macroeconomics. In this paper we have taken an explicit identification approach, and implemented it on data from the U.S. March Current Population Surveys (MCPS) covering earnings years 1963 to 2008. The approach, guided by standard human capital theory, is based on a modified version of the flat spot method proposed in Heckman, Lochner and Taber (1998). The price series all show substantial movement over the 1963 to 2008 period and the patterns are robust to a number of validity checks and sensitivity analysis. Perhaps the most surprising result is a very high correlation between the series from the lowest education group (high school dropouts) to the highest (college graduates). All of the series exhibit an increase in the price from 1963 to the mid-1970s followed by a substantial decline through the 1980s and 1990s that is interrupted by plateaus or recoveries coming out of the recessions of the early 1980s and the early 1990s. By the early 2000s the price series are all relatively stable.

The secular pattern of the price series has important implications for interpreting life cycle wage profiles. Using the standard alternative constant price assumption, the human capital profiles for different cohorts are very different in shape and show a confusing cohort pattern that is hard to interpret within a standard human capital model. By contrast, the use of the price series shows that the implied sequence of life-cycle human capital profiles for successive birth cohorts from the 1920s to the 1950s all have similar shapes and follow a simple pattern that is consistent with an optimal human capital model of the Ben-Porath type.

The high correlation between the price series across education groups implies that much of the change in relative wages by skill is due to changes in relative quantities rather than relative prices. We re-examined the college premium, and decomposed the change in relative wages into the separate components of relative price and quantity changes. This decomposition shows that relative quantity changes are at least as important as relative price changes in explaining the path of the average college premium. In addition, selection effects on relative quantities implied by cohort education patterns can also explain the age patterns, documented in Card and Lemieux (2001).

The maintained hypothesis explored in this paper is that college graduates have the same "type" of college graduate human capital over time. We contrast this hypothesis with an alternative skill biased technological change hypothesis under which there was a change in vintage "type", coinciding with the widespread introduction of micro computers, that resulted in an increase in the college premium for the young ("new vintage") college graduates, but not for the ("old vintage") older college graduates, as described in Card and Lemieux (2001). The primary difference between the two hypotheses is their different implications for the time path of vintage effects. Under the hypothesis explored in this paper vintage effects occur through selection effects and technological improvement in human capital production which occur continuously. The alternative hypothesis is that vintage effects come about through changes in vintage types rather than through quantity changes due to selection effects and technological improvement. Thus, for the maintained hypothesis of this paper, the time path of the vintage effects is constrained by the selection effects implied by cohort education patterns combined with technological improvement in the production of human capital. For the alternative hypothesis the pattern is constrained by the timing of the vintage type change.

Using estimates of the vintage effects across successive college graduate cohorts derived from estimated life cycle human capital profiles, the two alternative hypotheses may be compared in their fit with the data. The two hypotheses differ in their predicted patterns for the vintage effects. The estimates from the MCPS show results that are consistent with our maintained hypothesis where selection and technological progress in the production of human capital drive the vintage effects. In contrast, the estimates are inconsistent with the alternative hypothesis based on a vintage "type" change in the mid 1970s. The results suggest that a relatively parsimonious model can explain the wage patterns without the need for imperfect substitutability between college graduates of different ages or vintages.

Finally, the price series was used to construct a new quality adjusted measure of the labor input. For many issues of secular growth, cross country variation and cross cohort variation in wages, standard measures of human capital leave out important sources of progress or variation: technological change in human capital production, broadly interpreted, selection effects of changing educational attainment, and the increased life-cycle investment that accompanies the increased labor market attachment of women. Adjusting the labor input for quality changes using the estimated price series greatly reduces the contribution of MFP growth to recent growth in the standard of living in the U.S., and emphasizes the role of increases in the true labor input. This parallels some recent research suggesting that quality adjustment to international comparisons of human capital greatly reduces MFP differences as the source of cross country differences in wealth.[48] A large part of the increase in the quality of the labor input is not due to composition changes but instead to technological change in human capital production and changes in the optimal accumulation over the life-cycle, especially for females. Since most attempts at adjusting the labor input for quality changes, such as Krusell *et. al.* (2000) or the official BLS series used to estimate MFP, only deal with composition, they cannot capture a large part of the quality change.

References

[1] Gary S. Becker. *Human Capital.* National Bureau of Economic Research, New York, 1964.

[2] Yoram Ben-Porath. The production of Human Capital and the Lifecycle of Earnings. *Journal of Political Economy*, 75:352–365, 1967.

[3] Christopher Bollinger and Barry Hirsch. How Well are Earnings Measured in the Current Population Survey? Bias From Non-response and Proxy Respondents. mimeo, 2008.

[4] Audra J. Bowlus, Haoming Liu, and Chris Robinson. Human Capital, Productivity and Growth. CIBC Human Capital and Productivity Project Working Paper, 2005-2, University of Western Ontario, 2005.

[48] The most dramatic example is Manuelli and Seshadri (2005). Their estimates show very little cross country difference in TFP when the quality of human capital is taken into account. TFP in the poorest countries is not much smaller than that of the U.S. at around 73 percent of the U.S. figure. By contrast, studies that do not take into account human capital quality find rates for the poorest countries at only 20 percent of the U.S. value.

[5] Audra J. Bowlus and Chris Robinson. Human Capital Prices, Productivity and Growth. University of Western Ontario, CIBC Working paper No. 2008-5, 2008.

[6] Bureau of Labor Statistics, U.S. Department of Labor. *Labor Composition and U.S. Productivity Growth, 1948-90*, Bulletin 2426, 1993.

[7] Bureau of Labor Statistics, U.S. Department of Labor. *Handbook of Methods*, 1997.

[8] David Card and Thomas Lemieux. Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort Based Analysis. *Quarterly Journal of Economics*, CXVI:705–746, 2001.

[9] Peter Chinloy. Sources of Quality Change in Labor Input. *American Economic Review*, 70:108–19, 1980.

[10] Edward Denison. *Trends in American Economic Growth, 1929-1982*. Brookings Institution, Washington, DC, 1985.

[11] Jeremy Greenwood, Zvi Hercowitz, and Per Krusell. Long Run Implications of Investment Specific Technological Change. *American Economic Review*, 87:342–362, 1997.

[12] Fatih Guvenen and Burhanettin Kuruscu. A Quantitative Analysis of the Evolution of the U.S. Wage Distribution: 1970-2000. mimeo, 2007.

[13] Gary D. Hansen. The Cyclical and Secular Behavior of the Labor Input: Comparing Efficiency Units and Hours Worked. *Journal of Applied Econometrics*, 8:71–80, 1993.

[14] Gary D. Hansen and Selo Imrohoroglu. Business Cycle Fluctuations and the Life Cycle: How Important is On-the-Job Skill Accumulation? NBER Working Paper No. 13603, 2007.

[15] James J. Heckman. A Life-Cycle Model of Earning, Learning and Consumption. *Journal of Political Economy*, 84:S11–S44, 1976.

[16] James J. Heckman, Lance Lochner, and Christopher Taber. Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of labor Earnings with Heterogeneous Agents. *Review of Economic Dynamics*, 1:1–58, 1998.

[17] James J. Heckman, Lance Lochner, and Petra Todd. Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond. In Eric A. Hanushek and Finis Welch, editors, *Handbook of the Economics of Education*. Elsevier B.V., 2006.

[18] Barry Hirsch and Edward Schumacher. Match Bias in Wage gap Estimates Due to Earnings Imputation. *Journal of Labor Economics*, 22:689–722, 2004.

[19] Mark Huggett, Gustavo Ventura, and Amir Yaron. Human Capital and Earnings Distribution Dynamics. *Journal of Monetary Economics*, 53:265–290, 2006.

[20] Susumi Imai and Michael P. Keane. Intertemporal Labor Supply and Human Capital Accumulation. *International Economic Review*, 45(2):601–641, 2004.

[21] David Jaeger. Reconciling the Old and New Census Bureau Education Questions: Recommendations for Researchers. *Journal of Business and Economic Statistics*, 15:300–309, 1997.

[22] Matthew Johnson and Michael P. Keane. A Dynamic Equilibrium Model of the U.S. Wage Structure, 1968-1996. mimeo, 2008.

[23] Dale W. Jorgenson, Frank M. Gollop, and Barbara M. Fraumeni. *Productivity and U.S. Economic Growth*. Harvard University Press, Cambridge, MA, 1987.

[24] Lawrence Katz and David Autor. Inequality in the Labor Market. In Orley C. Ashenfelter and David Card, editors, *Handbook of Labor Economics*. North Holland, 1999.

[25] Lawrence Katz and Kevin Murphy. Changes in Relative wages, 1963-1987: Supply and Demand factors. *Quarterly Journal of Economics*, 107:35–78, 1992.

[26] Per Krusell, Lee E. Ohanian, Jose-Victor Rios-Rull, and Giovanni L. Violante. Capital Skill Complementarity and Inequality: A Macroeconomic Analysis. *Econometrica*, 68:1029–1053, 2000.

[27] Burhanettin Kuruscu. Training and Lifetime Income. *American Economic Review*, 96:832–846, 2006.

[28] Finn E. Kydland and Edward C. Prescott. Cyclical Movements of the Labor Input and Its Implicit Real Wage. *Federal Reserve Bank of Cleveland Economic Review*, 29:12–23, 1993.

[29] Donghoon Lee. An Estimable Dynamic General Equilibrium Model of Work, Schooling and Occupational Choice. *International Economic Review*, 46(1):1–34, 2005.

[30] Lee Lillard, James Smith, and Finis Welch. What do We Really Know about Wages? The Importance of Nonreporting and Census Imputation. *Journal of Political Economy*, 94:489–506, 1986.

[31] Rodolfo E. Manuelli and Ananth Seshadri. Human Capital and the Wealth of Nations. mimeo, 2005.

[32] Jacob Mincer. *Schooling, Experience and Earnings*. National Bureau of Economic Research, Columbia University Press, New York, 1974.

[33] Sherwin Rosen. A Theory of Life Earnings. *Journal of Political Economy*, 84:S45–S67, 1976.

[34] Robert Solow. Technical Progree and the Aggregate Production Function. In Charles R. Hulten, Edwin R. Dean, and Michael J. Harper, editors, *New Developments in Productivity Analysis*. University of Chicago Press, 2001.

[35] Yoram Weiss. The Determination of Life Cycle Earnings: A Survey. In Orley C. Ashenfelter and Richard Layard, editors, *Handbook of Labor Economics*. North Holland, 1986.

[36] Yoram Weiss and Lee Lillard. Experience, Vintage and Time Effects in the Growth of Earnings: American Scientists, 1960-1970. *Journal of Political Economicy*, 86:427–447, 1978.

A Appendix

The data for the analysis come from the March Current Population Surveys (MCPS). A consistent and annotated version of the files from UNICON was used as the data source. In this Appendix these data are described with particular reference to issues of data quality and comparability over time in Sections A1-3. Section A4 documents the robustness of the flat spot estimates presented in Figure 3, and Section A5 presents the alternative standard unit estimates for the dropouts.

A.1 Consistent Education Categories

The issue of consistency of the education measure arises because of a break in the education questions in 1991. This break is studied in detail in Jaeger (1997) who compared the education answers from the same respondents at different points in their CPS rotation who were asked the old education questions in their earlier rotation and the new questions in their later rotation. Jaeger offers solutions of two types. First is a linearization of the new educational attainment question that approximates the old "highest grade completed". The recommended mapping to construct a consistent "highest grade completed" or "years of schooling" variable is provided in the first and last columns of Jaeger's Table 2. Second, Jaeger considers 4 category matches rather than linearization. These are high school dropouts, 12th grade, some college, and college graduates. The recommended mapping for creating these four categories consistently across time is given in Jaeger's Table 6. In this paper we use the same four categories as Jaeger and follow his category mapping across the break.[49]

A.2 Consistent Annual Hours Measures

The MCPS annual labor incomes are for the year preceding the survey. Prior to the 1976 survey (1975 earnings) reported working hours in the survey could not be related to the previous year's earnings. In the MCPS data, hourly wages can be constructed as the ratio of annual labor income to annual working hours. Annual working hours can be constructed as the product of weeks worked per

[49] There is a small difference in this mapping from a standard high school dropout/high school graduate cutoff using the linearization. For the period 1975-1990 this is the same under both sets of coding. However, for the period 1991-2001, in contrast to mapping of code 38 into the less than high school group, Jaeger's category mapping puts them into high school. This is due to the use of the median rather than the mean in Table 2. The mean of the 38 group is actually 11.38 but the median is 12. Up to 1990 the fraction of high school dropouts is the same under both definitions. The cumulative fraction up to and including 11 over the $1985-1990$ period was 17.76, 17.36, 17.31, 16.86, 16.68 and 16.09. Jaeger's category mapping takes it to 13.40 for 1991 and the alternative takes it to 15.12. Further inspection, however, shows that the big drop is actually in the cumulative to 10 years which is common to both measures, so the less drastic drop from the alternative method is not to be preferred on this ground.

year and usual hours worked per week for the 1976 survey onward. Prior to this survey year, usual weekly working hours were not recorded and weeks worked were reported in grouped categories. An imputation procedure was used to create a series back to the 1964 survey.

- *Hours Worked per Week Last Year.* For the surveys before 1976, the MCPS variable *hrslyr* ("hours last year") is not available, and an estimate has to be obtained from data on *hours* ("hours last week"). The question for this variable is always the same: "In the weeks that ..... worked, how many hours did ..... usually work per week?" An estimate of hours worked per week last year for the survey years prior to 1976 is constructed as follows. First, for the individuals who were working last week, their "hours last week" is used as an estimate of their "hours per week last year."[50] Second, for the individuals who were not working last week but who had worked last year, their "predicted hours last week" is used as an estimate of "hours per week last year" where the predicted hours is obtained from a regression of "hours last week" on age, education in years and a female dummy variable for each year on the sample of those employed in the survey year.

- *Weeks Worked per Year Last Year.* For the 1962-1975 surveys the question is: "In 19XX how many weeks did ..... work either full time or part time (not counting work around the house) ?" For the 1976 and 1977 surveys the question was amended to: "In 19XX how many weeks did ..... work either full time or part time, not counting work around the house? Include paid vacation and paid sick leave." From the 1978 survey on, the question became: "During [19XX/20XX] (last year) in how how many weeks did ..... work even for a few hours? Include paid vacation and paid sick leave." Prior to the 1976 survey this variable was only available in intervals. UNICON created a time consistent variable for "weeks worked last year" by using interpolated values based on interval means from some post-1975 surveys.

A.3 Consistency and Quality Issues for the Annual Earnings Measure

The annual wage and salaries earnings data are from the UNICON time consistent "income from wage and salary" variable derived from the MCPS variable *incwag* (income from wage and salary). The definition from the glossary is as follows: *Money wages or salary is defined as total money earnings received for work performed as an employee during the income year. It includes wages,*

[50] The question for the hours variable in survey years 1962-1993 was: "How many hours did ..... work LAST WEEK at all jobs?"

salaries, Armed Forces pay, commissions, tips, piece-rate payments and cash bonuses earned, before deductions are made for bonds, pensions, union dues, etc. Earnings for self-employed incorporated businesses are considered wage and salary.

The question for the survey years 1963-1968 is "Last year how much did ..... receive: In wages or salary?" For the survey years 1969-1974, the question was slightly amended to "Last year ($19XX$) how much did ..... receive: In wages or salary before any deductions?" and for survey years 1975-1979 was further amended to "Last year ($19XX$) did ... receive any money in wages and salary? If so, how much did ... receive before any deductions?" From 1980 onwards there are multiple questions for the source so that "income from wage and salary" is a sum of components, but there is a single top-code variable that applies to the to the total. From $1988B$ to 1995 the construction is (*incer1* if *ernsrc=1*) + *incwg1*, where *incer1* is the CPS "income from the longest job", *ernsrc* is 1 if the source of income from the longest job is wage and salary, and *incwg1* is the CPS "income from other wage and salary." There are two top-code flags for this period, one for *incer1*, and one for *incwg1*, hence the "income from wage and salary" variable can have a value above any single top-code cut off value. While the form of the question has been relatively stable over time, several potential quality issues arise from substantial time variation in the incidence and treatment of top-coding, and in allocated values.

Top-coding

The top-coding flag for the total *incwag* was not introduced until 1976. For $1964-1967$ the highest value of *incwag* is 99900, but there is no apparent top-coding from inspection of the frequencies. For $1968-1975$ the highest value is 50000 and there is clear top-coding from the frequencies, though without a flag it is not possible to say which of the observations with value 50000 are top-coded. For the years 1976 to 1981 the highest value is 50000; the top-coded observations can be identified from the flag except for 1977 when the flag indicates far too many top-coded and must be incorrect; the frequency at 50000 for 1977 strongly suggests top-coding at 50000. (The conditional frequency of 50000, given that the observation is above 45000, is almost the same as 1976.) The annual frequencies of top-coding for 1976-1981 are: 0.24, (30.83), 0.36, 0.51, 0.68 and 0.94. For $1982-1984$ the highest value is 75000. It is possible to say which of the observations with value 75000 are top-coded from the top-coding flag; the information from the flag and the frequencies agree. The frequencies for these years are: 0.37, 0.47 and 0.52. For $1985-1988$ the highest value is 99999. It is possible to say which of the observations with value 99999 are top-coded from the flag, except for 1985 where the flag must be incorrect.[51] The frequencies are: (0), 0.42, 0.54 and 0.63.

Beginning in 1989 (1988B) top-coding is done separately on the two components of *incwag*: income from the longest job last year (*incer1*), and other wage and salary income (*incwg1*). For 1989-1995 the top-coded value for *incer1* is 99,999; the flags all appear to be correct. The top-coded value for *incwg1* is also specified as 99,999 for 1989 to 1995, However, there are problems with the flag. For 1989 the flag is present but all values are missing.[52] The frequencies of top-coding on *incer1* (for positive values of *incer1*) and on *incwg1* (for positive values of *incwg1*) are as follows:

Year	*incer1*	*incwg1*
1989	0.80	-
1990	1.08	0.00
1991	1.05	0.00
1992	1.08	0.00
1993	1.23	0.00
1994	1.54	0.15
1995	1.77	0.21

The incidence of top-coding doubles over this period, reaching close to 2%. For calculating the price series, especially for the flat spot age group of college educated workers, this is a potential concern since the incidence for this group of relatively high earners can be much higher. A greater concern is the break in treatment at 1995/96.

For 1996 to 2002, values above 150,000 on *incer1* and above 25,000 on *incwg1* are replaced by demographic cell averages. These averages are apparent from the frequency tabulations. The flags for 2000 are obviously incorrect. The replacement values for *incwg1* in 2000 also has the extreme value 236224 for 6 observations. For 2003 onwards, cut off values are raised to 200,000 for *incer1* and above 35,000 for *incwg1*; values above these are again replaced by demographic cell averages and these averages are again apparent from the frequency tabulations. The frequency of top-coding and the most frequent replacement values are as follows:

[51] For 1985 all values of the flag are zero (no top-coding), despite a mass point at 99999 similar to adjacent years where the flag indicates top-coding.

[52] Unicon Appendix H4 notes some general problems with component top code flags, but apart from the problem with 1989, the frequencies are generally consistent with extremely low levels of top coding on *incwg1* throughout, so the pre-1994 zeros could be true. The 99,999 cut off is high for this "other wage and salary" component and was subsequently reduced.

Year	*incer1*	*incwg1*	replacement *incer1*	replacement *incwg1*
1996	0.60	2.62	302539	64524
1997	0.71	2.12	318982	45749
1998	0.79	3.84	330659	61345
1999	0.83	3.69	306731	59925
2000	(89.30)	(96.88)	(229339)	(50037)
2001	1.07	5.58	335115	56879
2002	1.25	5.31	320718	60670
2003	0.74	2.66	390823	91360
2004	0.72	3.27	404469	89988
2005	0.71	3.27	422850	77282
2006	0.81	3.54	423545	79378
2007	0.92	4.21	437528	74091
2008	0.90	4.33	419969	73029
2009	1.05	4.51	389599	72946

The frequency of top-coding for each component changes substantially between 1995 and 1996 with the changes in top-coding cut-offs; the increase for *incer1* cuts the incidence on that component by two thirds, while the decrease for *incwg1* increases the incidence on that component more than 10 fold. The top-coding on the *incwg1* ("other wage and salary") varies substantially and reaches over 5% in some years. While the top-coding incidence changes are a concern, the shift to replacement values has the most dramatic effect. In general this effect is apparent in the upper tail, but the change is large enough to be clearly apparent in the mean wage for the whole sample. This is shown in the following table that reports the mean, median and maximum wage for *incwag*, as well as the 90th and 99th percentile.

Year	Mean	Median	90th Percentile	99th Percentile	Maximum
1990	24903.5	21000.0	50000.0	99999.0	199998.0
1991	25332.1	21223.0	50000.0	99999.0	180000.0
1992	25770.8	22000.0	51000.0	99999.0	199998.0
1993	26759.1	22428.0	54000.0	99999.0	193999.0
1994	27738.5	23180.0	56732.0	99999.0	199998.0
1995	29290.7	24648.0	60000	99999.0	199998.0
1996	32143.2	25000.0	60000.0	257390.0	464782.0
1997	33189.9	25000.0	61500.0	318982.0	454816.0
1998	35098.2	26999.0	66000.0	330659.0	418608.0
1999	36767.8	28831.0	70000.0	306731.0	492657.0
2000	36975.6	29000.0	72000.0	229339.0	364302.0

In 1995, the last year of non-replacement top-coding, the highest value is 199998, the 99th percentile is 99999, the 90th percentile is 60000, the mean is 29291 and the median is 24648. In

1996, the highest value jumps to 464782, the 99th percentile jumps to 257390 and the mean jumps to 32143, while the 90th percentile and the median both show modest or no increase in line with previous years and subsequent years.

Two additional concerns are the effect of the problem flags for 2000 and the extremely large value of 240674 for 8 observations used as the replacement values for *incwg1* in 2007. The effect of the problem flags for 2000 is shown in the following table that reports the top coding counts for *incer1* and *incwg1* and the mean income conditional on being above or at least equal to $149000

Year	top-coding count *incer1*	top-coding count *incwg1*	mean > 149000	mean $<= 149000$
1996	372 (0.58)	309 (0.48)	274678.2	23587.27
1997	427 (0.65)	227 (0.35)	296586.4	24249.76
1998	471 (.073)	451 (0.70)	296775.1	25548.10
1999	511 (0.78)	428 (0.65)	297244.1	26808.38
2000	59543 (88.89)	64705 (96.59)	219548.8	27546.18
2001	659 (1.02)	577 (0.89)	294638.7	28954.72
2002	1304 (1.23)	852 (0.80)	292384.1	29989.74
2003	774 (0.74)	381 (0.36)	279471.4	30647.21
2004	716 (0.70)	438 (0.43)	263159.5	31345.67
2005	711 (0.71)	436 (0.43)	273961.5	32092.19

Clearly, the flag problem with 2000 results in a large change in the upper tail for that year. The mean, conditional on *incwag* $>$ 149000 shows a smooth progression over the years, including for 2000. The mean conditional on *incwag* $<=$ 149000 shows an abrupt fall for 2000.

Allocated Values

Allocated values are a serious issue in the MCPS data both because in some years as many as 25 percent of the values may be allocated, and because of time varying procedure for assigning allocated values.[53] The main change in treatment happened after the 1988 survey when the entire supplement was evaluated for response quality and the supplement information was deemed either a "good match" to the basic record or not. If the supplement was deemed a good match, the allocation procedure for the supplement information was the same as for the basic record with some variables being subject to having values allocated, as indicated by an allocation flag. If the supplement was not deemed a good match, then the entire supplement was allocated. The fractions allocated for the

[53]The allocation and nonresponse problem was discussed at length in Lillard, Smith and Welch (1986). More recently Bollinger and Hirsch (2008) drew attention to the serious problem of proxy responses and allocated values in CPS data. Hirsch and Schumacher (2004) document a dramatic example of how very misleading results can be obtained without careful treatment of the allocated values

income variable prior to 1989 ranged from around 11-18%. From 1989 there was a steady increase in the fraction of allocated values from around 18% in 1989 to over 30% in the mid 2000s. The allocation flag for income after 1988, only indicates an allocated value applied to a good match and only accounts for a minority of the allocations. For example, in 1989 there are 71226 records with positive *incwag*. Of these, 6963 were allocated as a results of the entire supplement being allocated. Then an additional 5504 were allocated as a result of the *incer1* part of a "good match" supplement being allocated, and these received an income allocation flag.[54]

A.4 Robustness of the Flat Spot Estimates

The main sample used in the paper is the FTFY sample; the preferred wage measure is the median wage. The results in the paper, however, are robust to alternative choices of samples and measures. The choice of sample and the wage measure are connected in that the use of medians largely avoids the problems of including or excluding top-coded or allocated values. Top coding is negligible for the flat spot samples for dropouts and high school graduates, and very low for some college.[55] However, top-coding is very important for the college graduate flat spot sample where the rates are highly variable and can reach close to 10%.

Using the full sample with all allocated values and no corrections for top-coding changes, income from wage and salaries (*incwag*) shows an obvious break at the major point of top-coding changes. The real hourly wage shows the same break in 1995/96, but in addition shows a number of other large jumps in the average wage relative to the median and the 90th percentile due to outliers. Most of the really major outlier problems are removed by the mild restriction of requiring at least 5 weeks of work for at least 5 hours per week. This drops only 1.6% of the sample. There do remain some very large hourly wage rates, but these are removed if the sample is further restricted to full-time and full-year (FTFY) workers, defined as working at least 40 weeks a years for at least 35 hours per week. The basic results are insensitive to the alternative sample restrictions on hours and weeks worked. This is illustrated in Figures A1a & A1b comparing the price series using the minimally restricted sample (requiring at least 5 weeks of work for at least 5 hours per week) instead of the

[54]The allocation problem is most serious for the income variables, but after 1988 when the entire supplement was allocated for many records, the hours and weeks last year were also subject to substantial allocation. However, compared to the income variable, only a small number of hours or weeks are allocated and flagged in the "good match" supplements.

[55]For dropouts, 1997 and 1998 have the highest rates, just below 0.5%, but for most years top coding is negligible. For high school, the rates are a little higher, reaching almost 1% in 2002, but for most years rates are below 0.5%. Some college has slighly higher rates, this time with several years around 1.5%, though most years are a lot smaller.

FTFY sample for high school graduates and college graduates, respectively.

Figure 3 used all the FTFY observations, including top-coded and allocated values. As noted earlier, the treatment of allocated values changed over time and could have an effect on the results. The fraction allocated, especially for more recent years, is much higher for the higher earning college graduates. Figures A2a and A2b show the sensitivity of the series to inclusion or exclusion of allocated values. The high school graduate series is virtually unaffected. The college graduate series is more sensitive, but the basic pattern with and without excluded values is the same. The same insensitivity to exclusion or inclusion of allocated values is true for log wages in place of median wages.

The flat spot samples used to construct the price series are particularly vulnerable to differences across pairs of years in the number (or treatment) of top-coded observations in the sample, especially for higher earners such as college graduates in their fifties. For example, in the period prior to the use of replacement values, over years when the nominal top-coding cutoff was constant, aging a cohort of college graduates is likely to cause a downward bias as an increasing fraction are subject to the cutoff. Conversely, when the cutoffs are abruptly increased, there may be an upward bias. In these examples, the use of medians avoids the bias problems.[56] More importantly, the switch to the use of assigning "average" wages among top-coded individuals of a given type instead of the top-coded (truncated) values is likely to create very serious bias problems, given the magnitude of the effect of this switch in treatment on mean wages for college graduates in their fifties. As shown earlier, the shift was important enough to have a significant effect on mean income for the whole sample. The effect was much more significant for college graduates in their fifties. The median and 90th percentile values of *incwag* were largely unchanged over the switch to replacement values between 1995 and 1996, whereas the mean shifted up almost 15%. This is directly reflected in a large shift up in the price series at this break if wages are used without taking into account this break.

The price series were all estimated with and without including top coded values. The series using medians are largely unaffected for all education groups. The series for the education groups with little top-coding are also insensitive to inclusion of top-coded observations, whether medians, average wages or average log wages are used. However, the series for college graduates which are most affected by top-coding are very sensitive to the inclusion of top-coded observations with a major

[56] Calculations performed in Bowlus and Robinson (2008) with and without the top-coded and allocated observations revealed potentially large biases when raw wages are used, as illustrated by a spike in their Figure A1 where an outlier in the MCPS data turns out to be one of the mean income replacement values assigned to a top coded observation.

break, as expected, at the shift to replacement values in 1995/96. The use of average wages shows the highest sensitivity. Only the series based on medians are insensitive to the treatment of top-coding. Ideally, the top coded observations should not be dropped, but in practice it appears difficult to include them without serious bias unless medians are used. In Figure A3 shows the price series using medians, excluding the top-coded observations. This is almost identical to Figure 3. Figures A4a & A4b show the similarity of the price series for high school graduates and college graduates based on three alternative underlying wage measures used to construct the annual differences for the flat spot groups: median wages, average wages, and average log wages. All measures use the same sample, excluding top-coded observations. The series for college graduates are slightly more sensitive in the in the period after 1995, but overall the general picture of Figure 3, and the high correlation of the series for the different education groups is robust to the alternative wage measures.

A.5 Robustness for the Dropout Series: Evidence from the Standard Unit Method

The standard unit method works by finding an observable "standard unit" of human capital that is the same across time. In this case, observing the wage paid for a standard unit at different points in time identifies the price change. This is similar to the notion of finding a time invariant common unit for computers. The solution in the computer case is to assume that the common unit that represents the factor provided by all computers is calculations per second. That is, calculations per second are the efficiency units. The relevant price is the price of a standard computer defined as having a given number of calculations per second.[57] Given the assumption of the common unit, the identification problem in the computer case is made very simple by the fact that computations per second can be observed so it is not necessary to actually observe "standard" computers over time to identify the relevant price. In the human capital case it is necessary to observe a standard unit over time because efficiency units are not directly observed. Implicitly, the standard unit approach is used in "composition bias" studies over the business cycle.[58] For studies of secular trends over longer time periods where quantities of human capital can change for any observable group, the choice of standard unit is more critical.

[57] Of course it is a little more complicated than this, since there are other dimensions on which computers may differ, and a hedonic analysis is often performed, but the basic idea is that a meaningful comparison can be made that permits an aggregation in terms of a standard unit.

[58] A regression approach to composition bias correction implicitly estimates a price series for the "omitted group" in a dummy variable framework.

The standard unit method uses the same estimating equation as the flat spot method,

$$ln\lambda_{t+1} = ln\lambda_t + D_{t+1},$$

where D_{t+1} is again the the difference in log wages between t and $t+1$ for a sample of observations which have the same efficiency units in t and $t+1$. But whereas the empirical counterpart to the D_{t+1} series for the flat spot method is obtained from following individuals in the same cohort over a period where their human capital does not change, the standard unit method replaces this by following a standard unit group across cohorts. The ideal standard unit group would be those members from the same lower tail portion of the initial endowment distribution that choose a zero level of further human capital production. By definition this group would have the same median human capital level across successive cohorts as well as across time. In practice, there is no such group. Instead it was approximated by a group where the addition to the initial endowment is the smallest, so that the human capital stock for this group is closely proxied by the initial endowment.[59]

While the main objective is to find a group that has the least contact with human capital production functions that may have been subject to technological change, it is also necessary to choose a group that has completed their education and become attached to the labor market, and a group such that the sample size is sufficiently large to yield reasonably precise estimates. It is also necessary to decide when to take the measure of completed education. Checks on the frequency distribution for experience by schooling group show that if individuals under 19 are included, the contemporaneously measured schooling completed for the lowest schooling group is not the correct final frequency - i.e. many go on to more education. By 19-20, however, those contemporaneously reporting a completed level of high school dropout correspond closely to the fraction that would report that same level at later ages. Thus, estimates using the standard unit method were restricted to samples aged 19 and above.

Finally, it is also necessary that the chosen group will have negligible selection effects over the time period of the earnings data, i.e. that successive cohorts will be drawn from the same tail of the initial endowment distribution. Figure 1 shows that for the earliest cohorts about a third of the cohort were high school dropouts. This was followed by a rapid decline until the first post-war birth

[59]The youngest dropouts have the least addition to their initial endowment, but they may also not be fully attached to the labor market. More importantly, this process may vary over time as different policies have been in place to help youth in training and transition to the labor market. There are also sample size considerations. Price series were estimated for several young age groups, producing similar results.

cohort when the fraction stabilized at around 13 percent. The earliest cohorts in the sample of 19-21 years of age are the 1942 to 1945 birth cohorts. Thus, apart from the first few years, the sample is obtained from successive cohorts with the same fraction of high school dropouts as required.

The benchmark flat-spot series in Figure 3 uses the FTFY sample. Examination of participation rates for the flat spot samples shows relatively constant rates across the period for all groups, including the dropout sample, for either the minimally restricted sample or the FTFY sample. The standard unit sample, however, is a much younger age group compared with the flat spot sample (19-21 vs. 44-53) and for them, while the minimally restricted sample shows approximately constant and relatively high participation, the FTFY restriction produces highly variable participation which is often less than 50%. Hence the FTFY is not suitable for standard unit. Figure A5 compares the benchmark flat spot series with standard unit estimates based on dropouts aged 19-21. Since the standard unit series uses the minimally restricted sample, the same sample is used for the flat spot estimates. The results show a close correspondence between the two methods except for the different recovery pattern out of the 1980 recession. The standard unit group showed more difficulty coming out of the recession, but eventually the series meet again.[60] Since the two methods are independent, this provides a partial check on the accuracy of the flat spot series for the dropouts.

[60] This may be due to participation differences.

TABLE 1

Cross Sections 1996-2007: Linear Slope Estimate over Ages 45-58

	Includes Allocated Values		Excludes Allocated Values	
	1	2	3	4
Year	Median	Log	Median	Log
1996	.0081	.0087	.0111	.0070
1997	.0098	.0038	.0068	.0021
1998	.0113	.0131	.0119	.0133
1999	.0080	.0094	.0068	.0090
2000	.0067	.0069	.0095	.0117
2001	.0027	-.0028	.0029	-.0014
2002	.0003	-.0026	-.0006	-.0036
2003	-.0014	-.0004	-.0039	-.0033
2004	-.0049	-.0054	-.0063	-.0063
2005	-.0102	-.0138	-.0100	-.0111
2006	-.0033	-.0074	-.0023	-.0057
2007	-.0033	-.0042	-.0062	-.0066

Note: Full-time, full year sample. Columns 1 & 3 report the slope using the log of the median wage as the dependent variable; columns 2 & 4 use the mean log wage.

TABLE 2

Estimated Life-cycle Human Capital Profiles by Cohort

Dependent Variable: Log Efficiency Units	
Birth Cohort	
1937	.0447
1940	.0420
1943	.0291
1949	-.0031
1952	.0178
1955	.0483
1958	.0865
1961	.1076
Age	.0914
Age squared	-.0009
Intercept	.5496
R^2	.8866

TABLE 3
Comparison of Alternative Labor Input Series, Paid Workers: 1977-2000

	Unadjusted Hours		Composition Adjusted Hours	Efficiency Units
	Total	Index	(Tornqvist)	
1977	146.2455	106.5429	106.5563	108.8724
1978	152.5765	111.1551	111.1707	113.5803
1979	155.6163	113.3697	113.3093	116.5712
1980	160.1237	116.6534	116.8424	118.2704
1981	161.009	117.2984	118.1511	119.2866
1982	158.9353	115.7877	117.4353	118.2980
1983	163.7512	119.2961	121.3482	120.0204
1984	172.2239	125.4687	128.1364	128.3710
1985	178.1313	129.7723	133.1505	133.0054
1986	182.9356	133.2724	137.3454	137.6373
1987	187.0298	136.2551	140.737	142.6102
1988	191.5991	139.5839	145.2336	150.3345
1989	195.0625	142.1071	148.6597	153.3596
1990	195.785	142.6334	149.7317	155.6080
1991	197.1212	143.6069	152.1546	157.3920
1992	198.7658	144.805	155.289	159.8216
1993	202.3861	147.4425	159.1748	169.2288
1994	208.1018	151.6065	165.1147	174.4911
1995	213.127	155.2675	169.105	180.8090
1996	215.9565	157.3288	171.8405	187.2434
1997	221.2122	161.1577	177.0715	192.1510
1998	227.0704	165.4255	183.4849	200.3117
1999	231.0326	168.3121	187.553	208.7188
2000	235.3721	171.4735	191.6503	219.3644

TABLE 4
Comparison of the Growth Rates of Alternative U.S. Labor Input Series:
Percentage Change 1977-2000

	Paid Workers 20-64		Private Sector Paid Workers 20-64	
EUS	101.49		111.63	
Hours	60.94		67.72	
BLS	79.86		90.30	
Fixed Weight	80.21			
	Males	Females	Males	Females
EUS	75.00	172.17	84.38	193.48
Hours	42.75	90.98	49.50	100.10
BLS (A)	56.21	142.96	65.80	163.91
BLS (B)	63.24	122.01	73.81	137.68

Figure 1

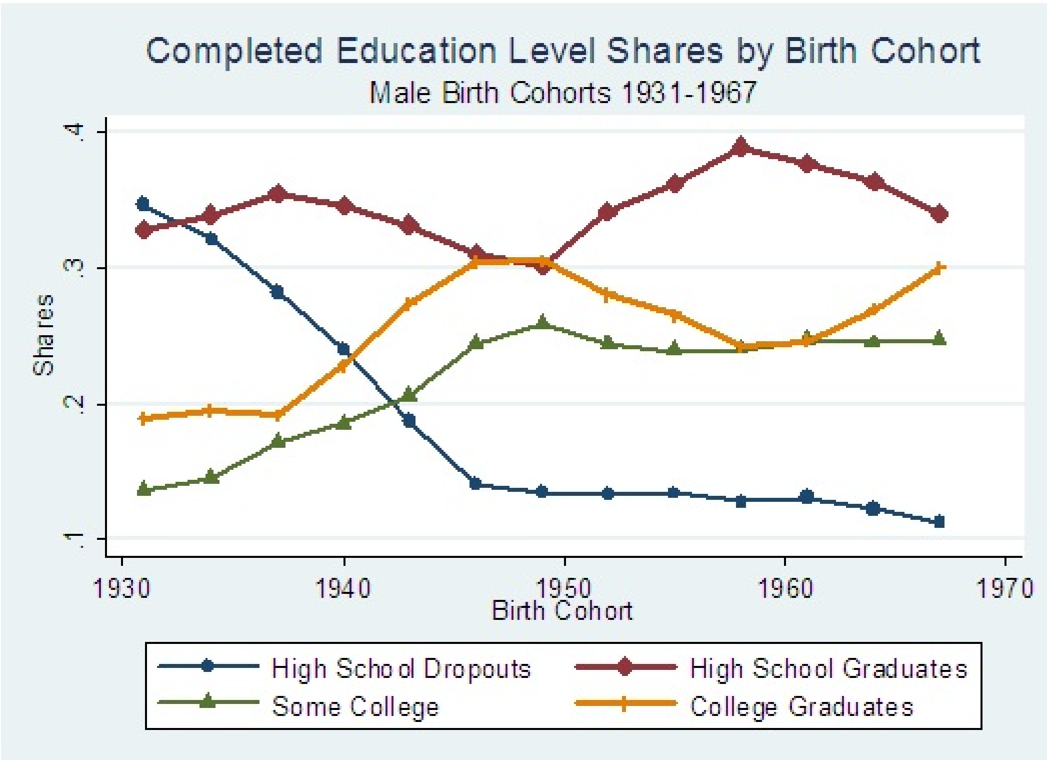

Figure 2a

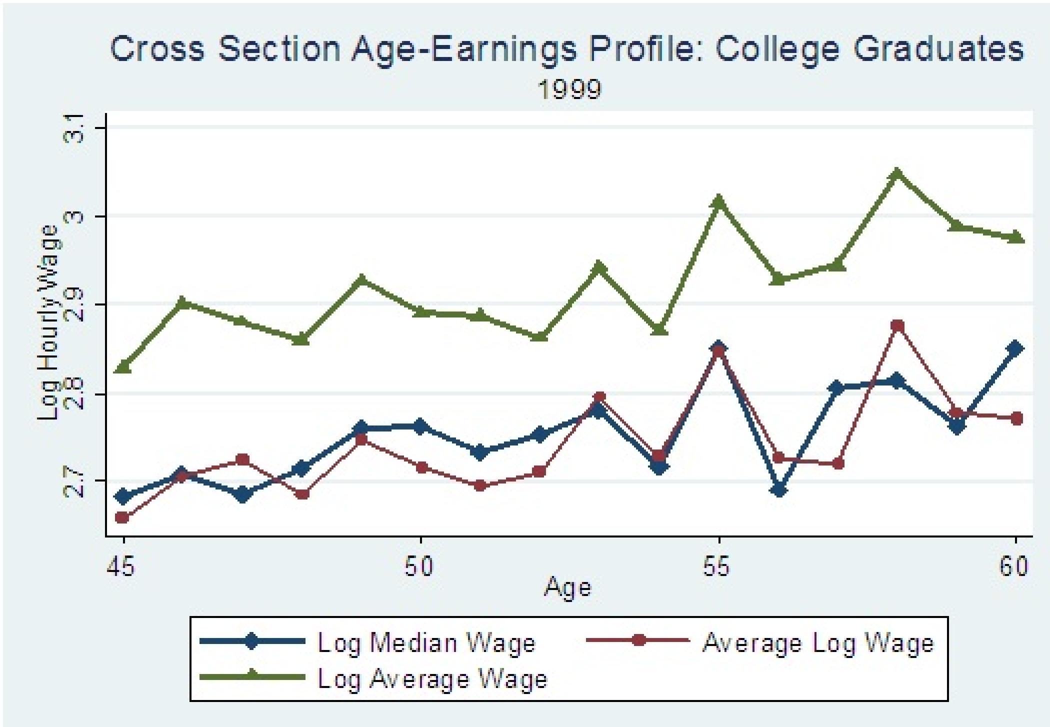

Figure 2b

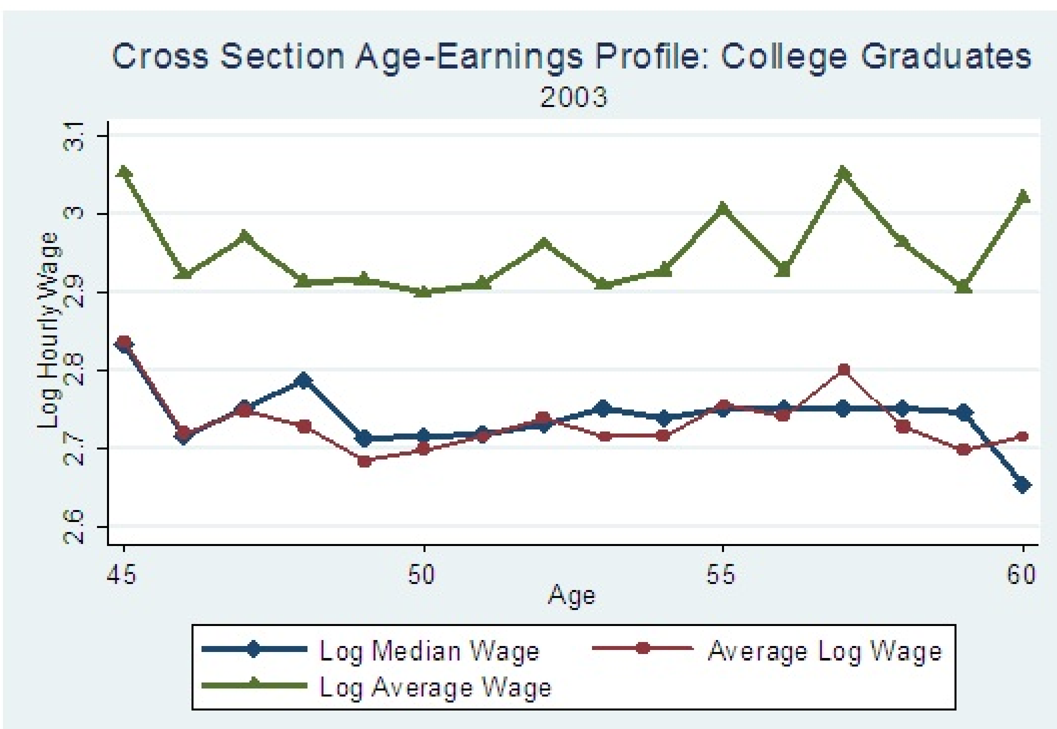

Figure 3

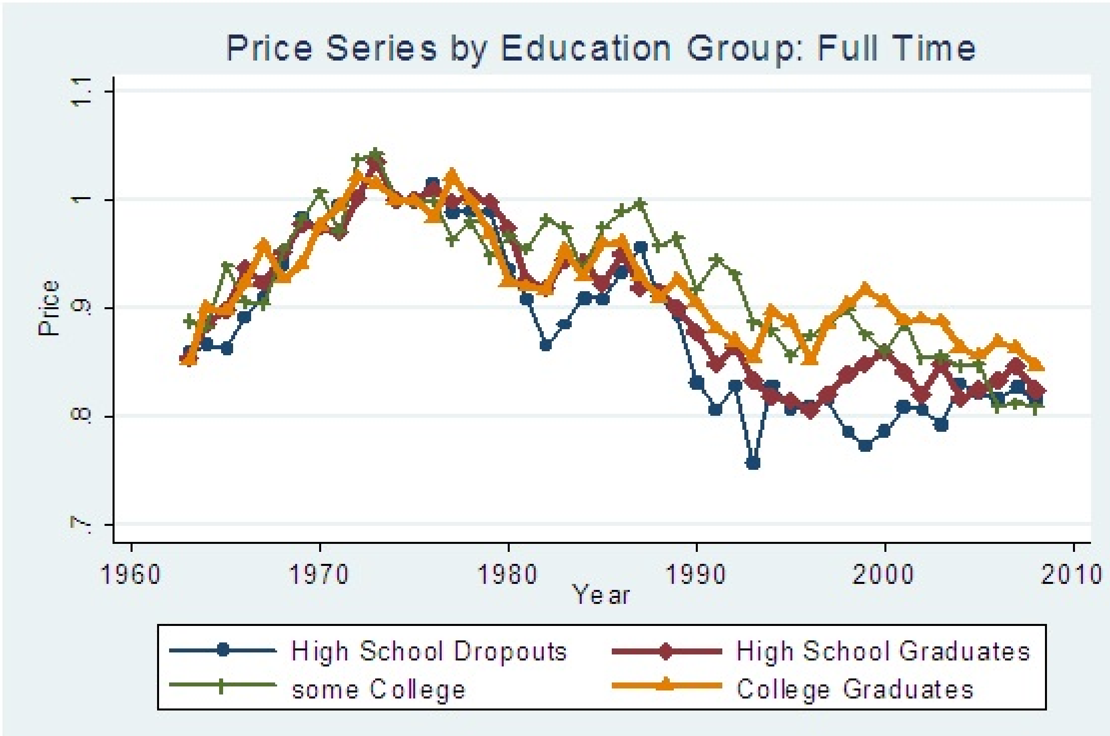

Figure 4

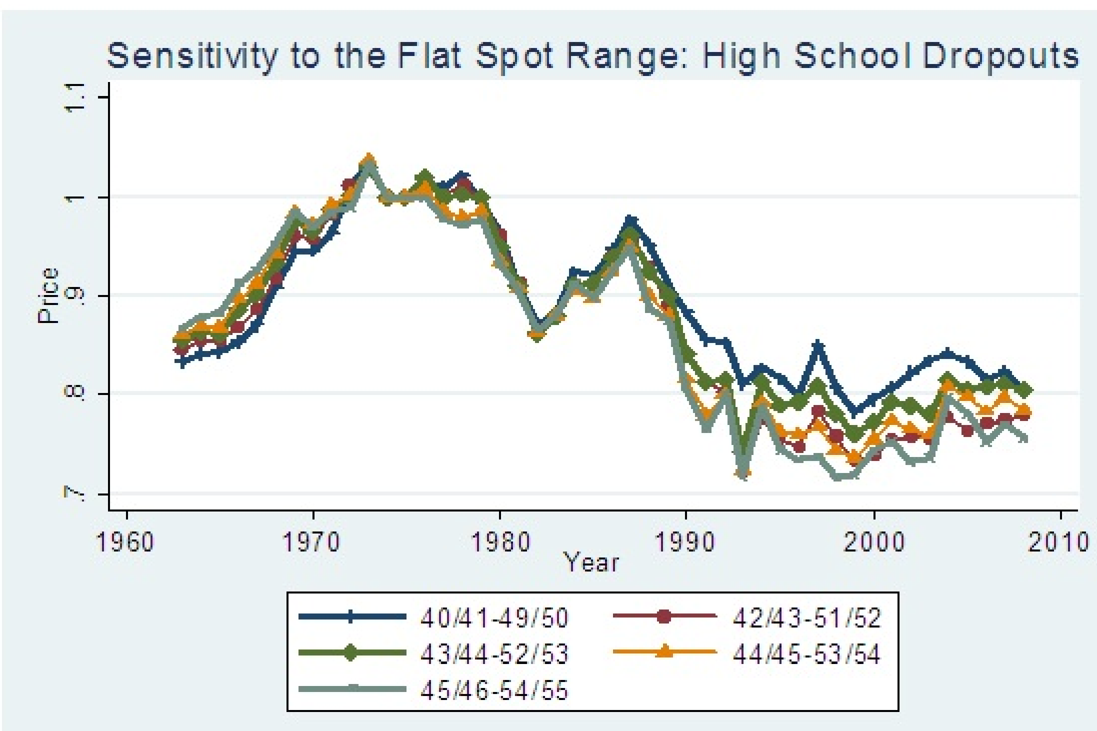

Figure 5

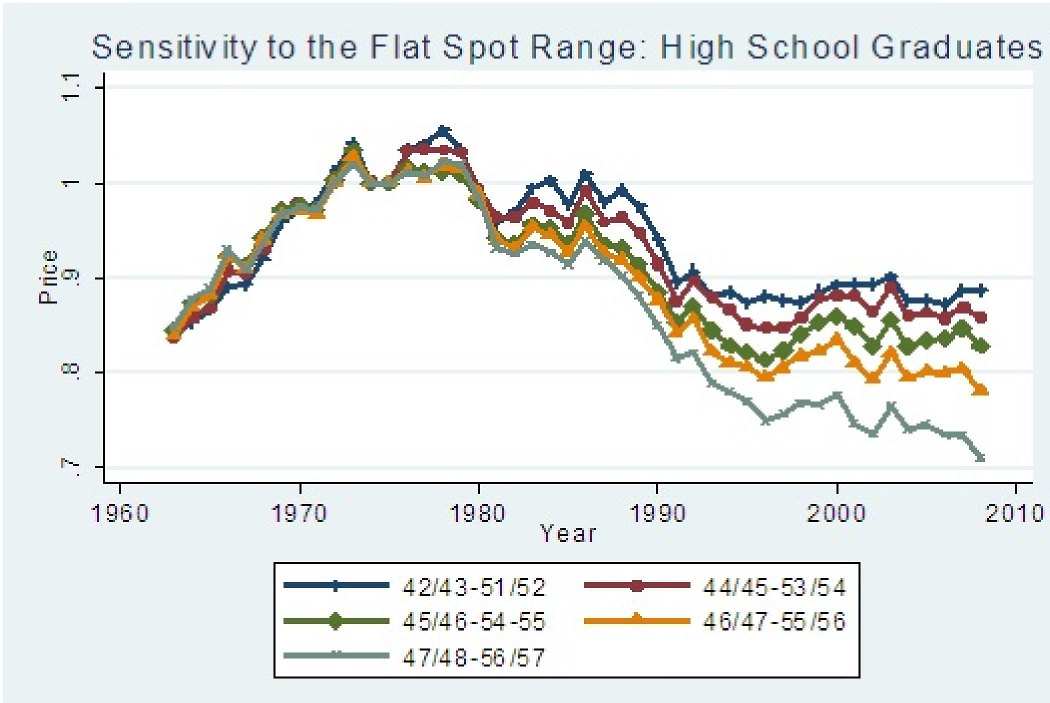

Figure 6

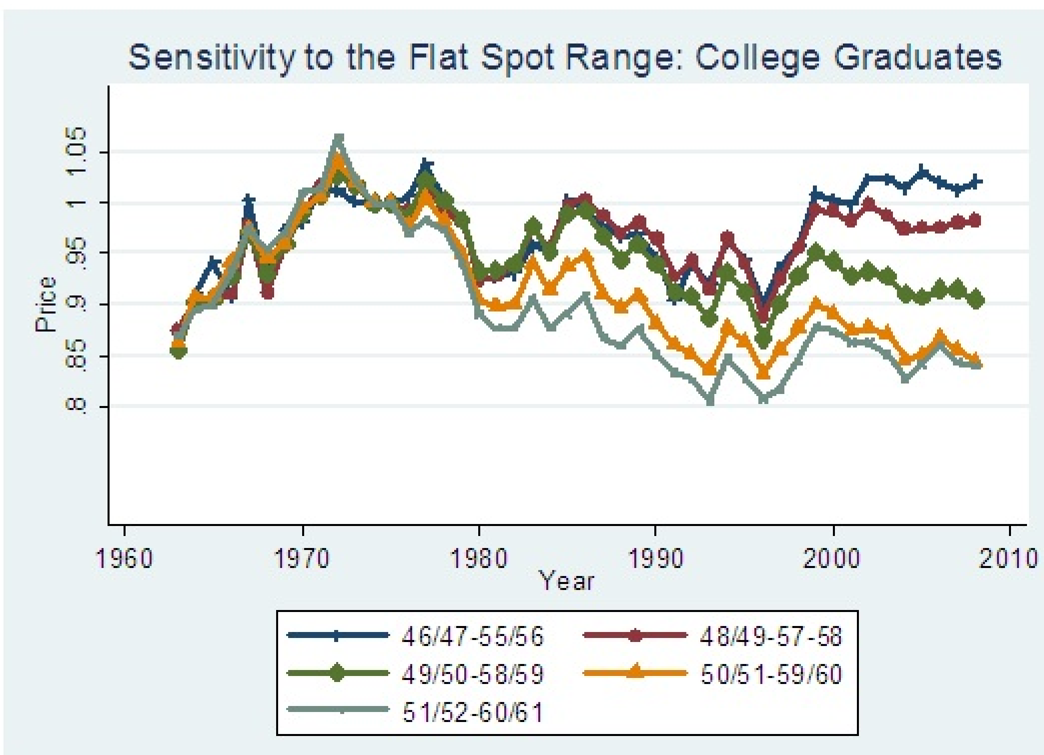

Figure 7

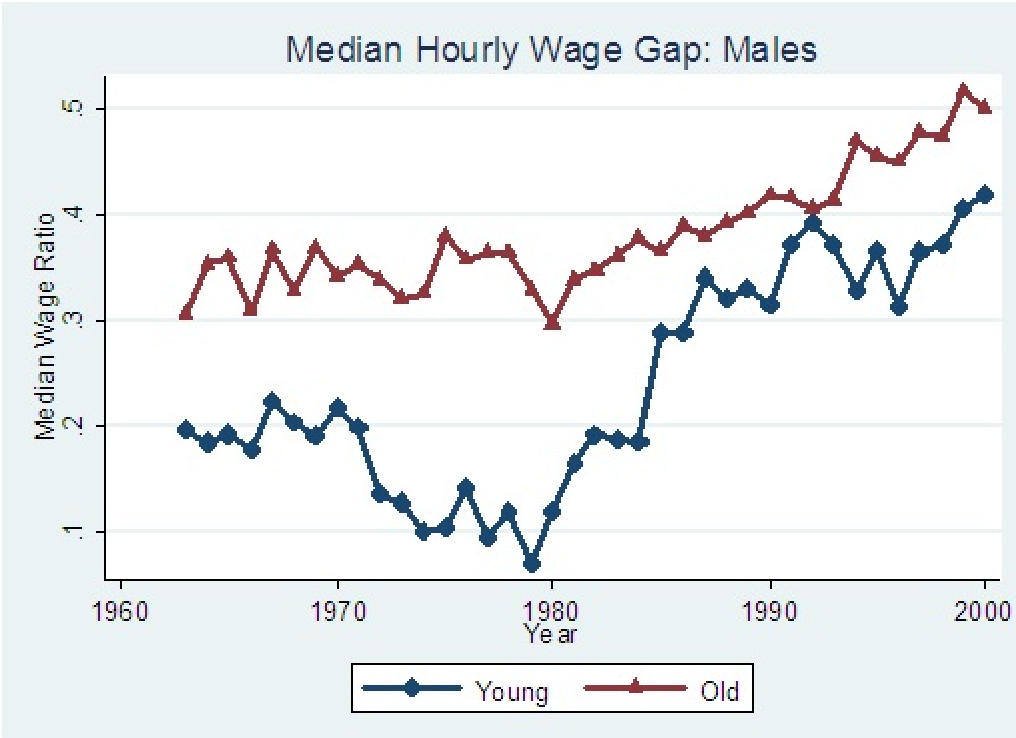

Figure 8

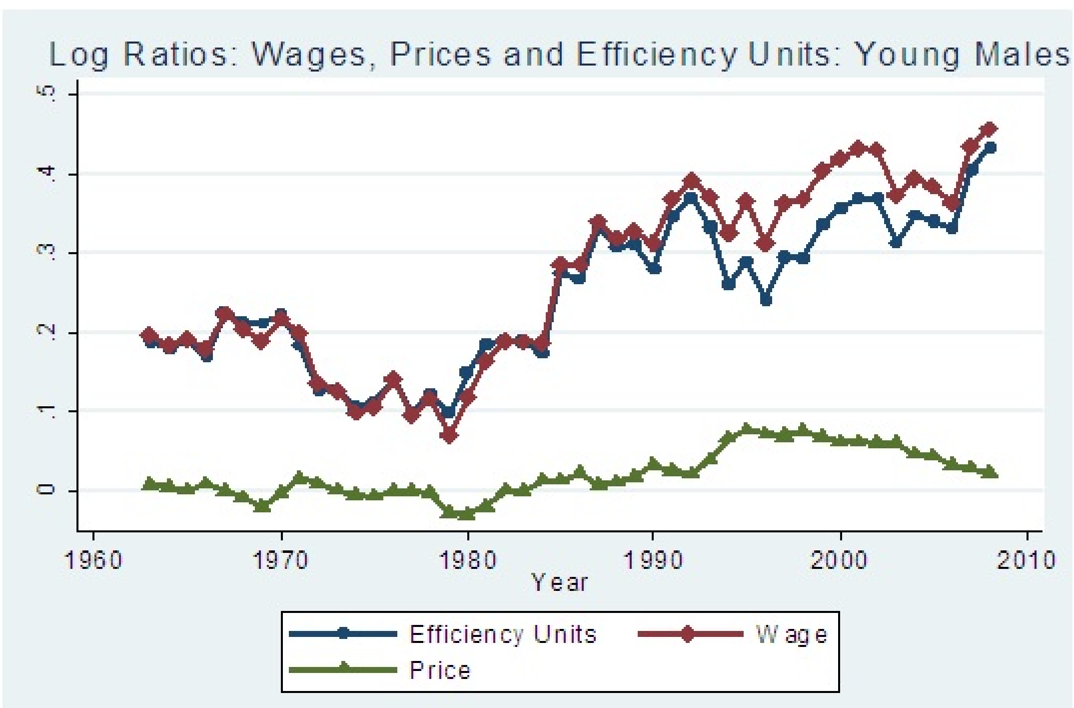

Figure 9a

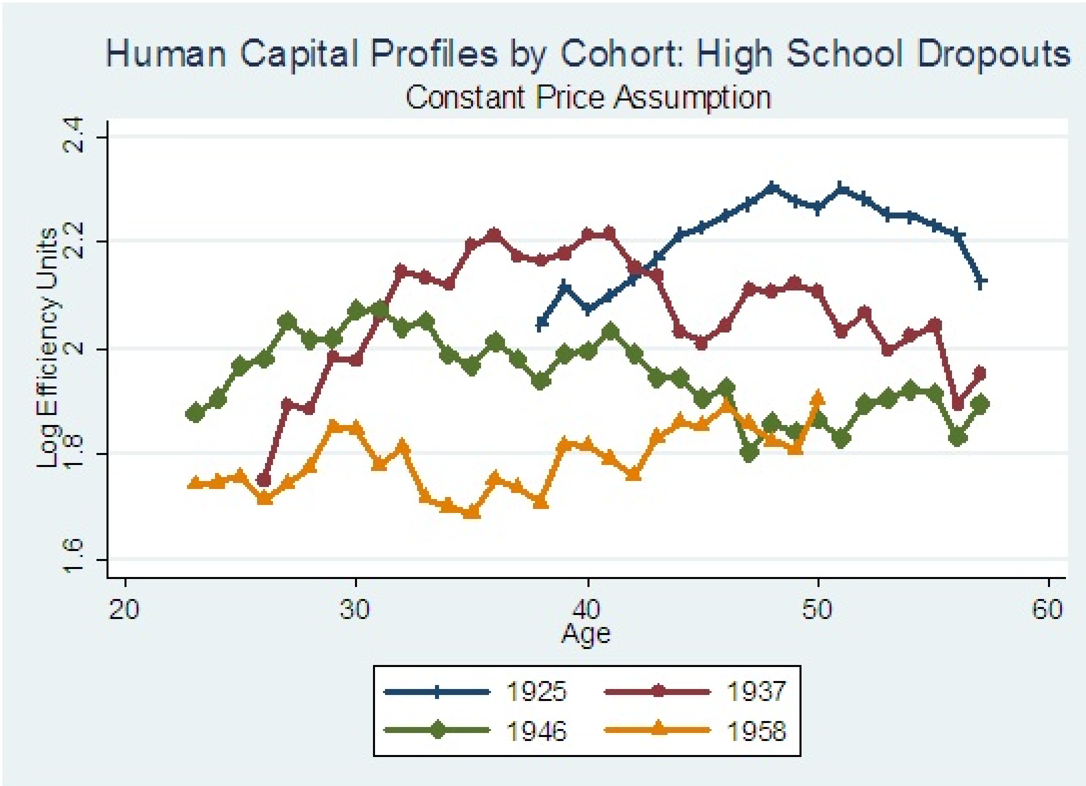

Figure 9b

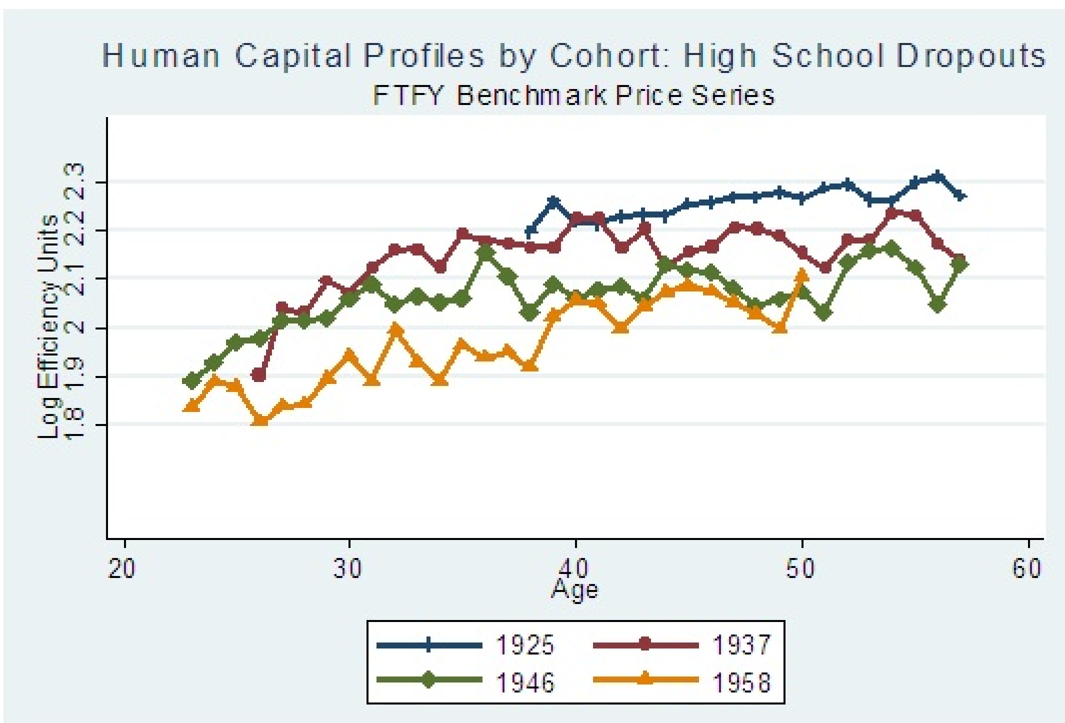

Figure 10a

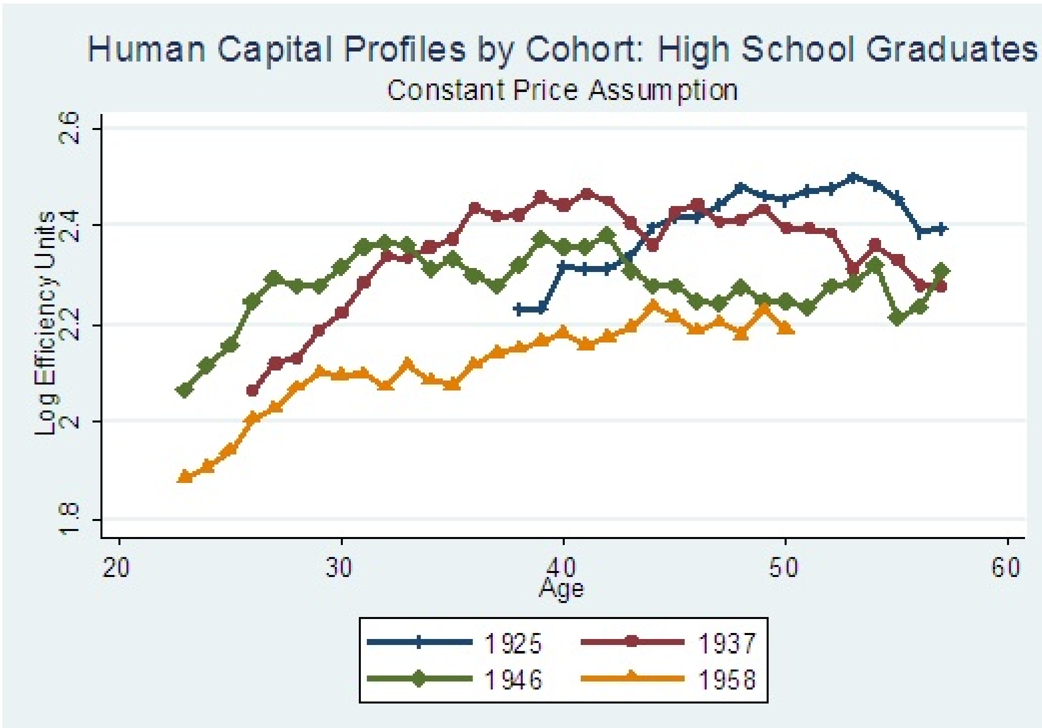

Figure 10b

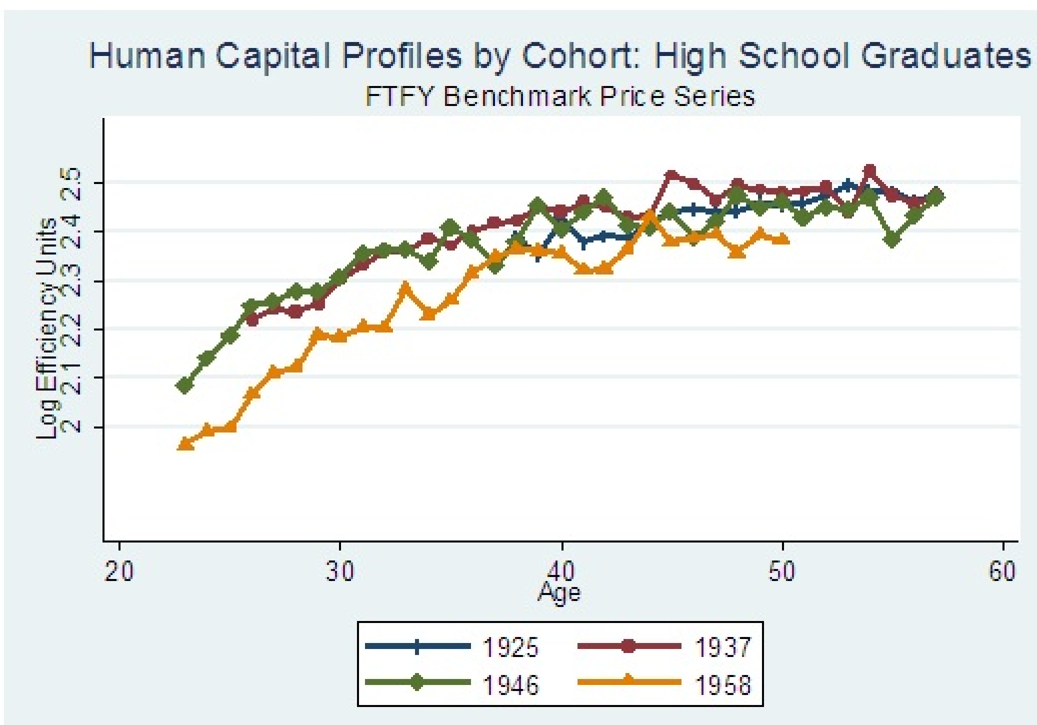

Figure 11a

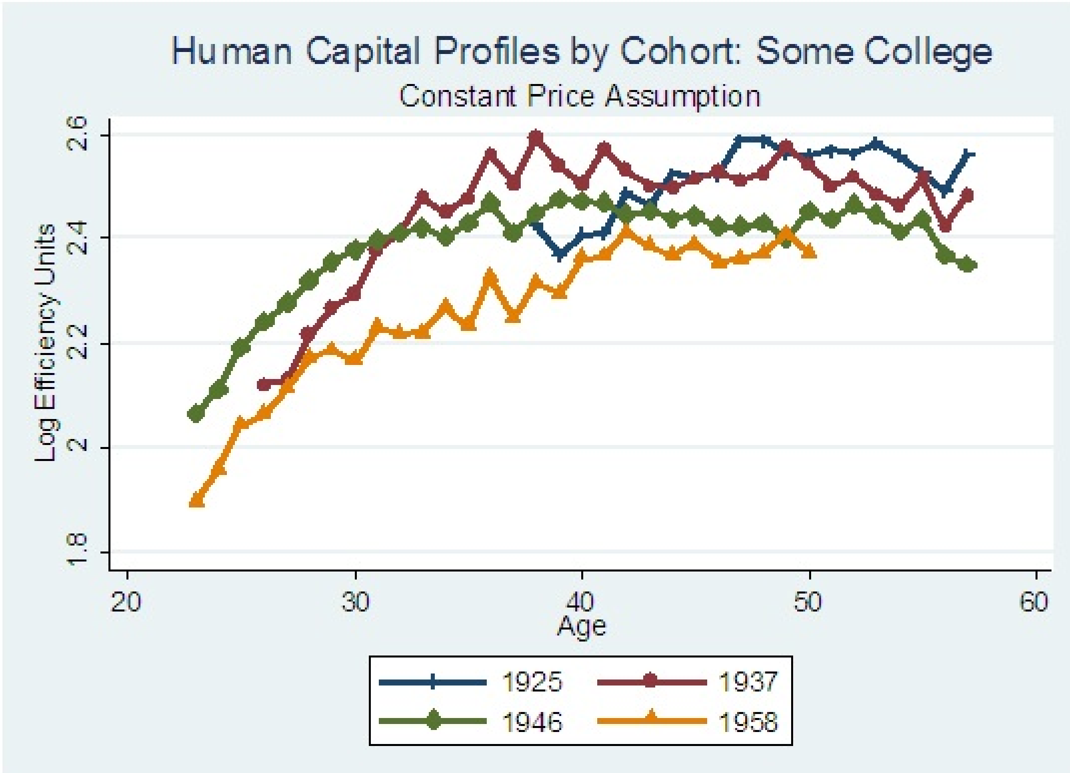

Figure 11b

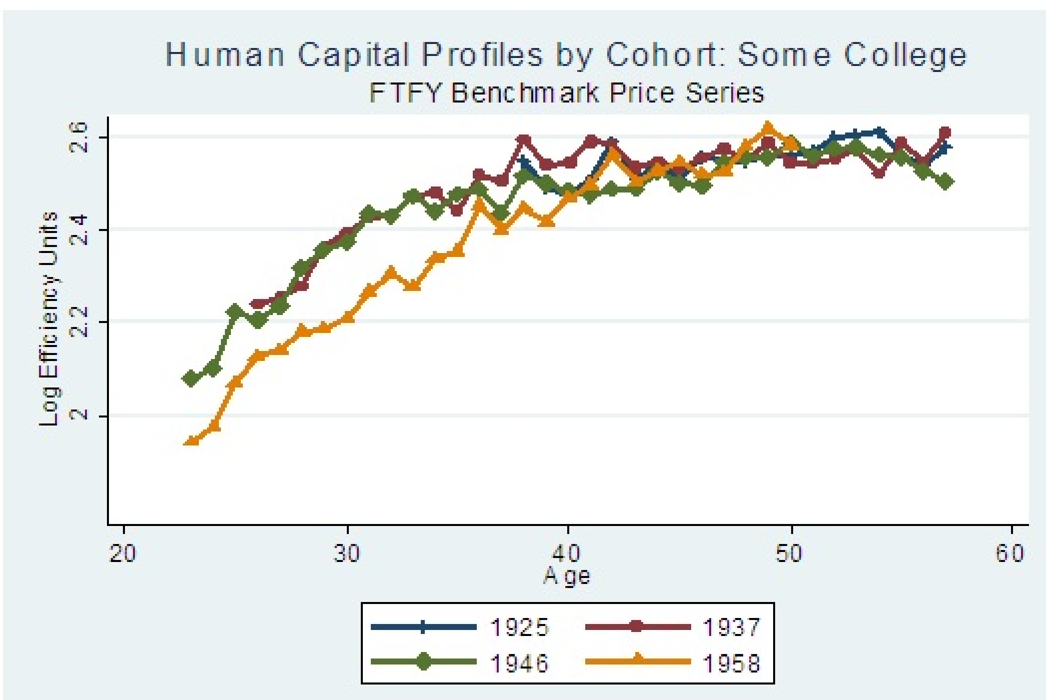

Figure 12a

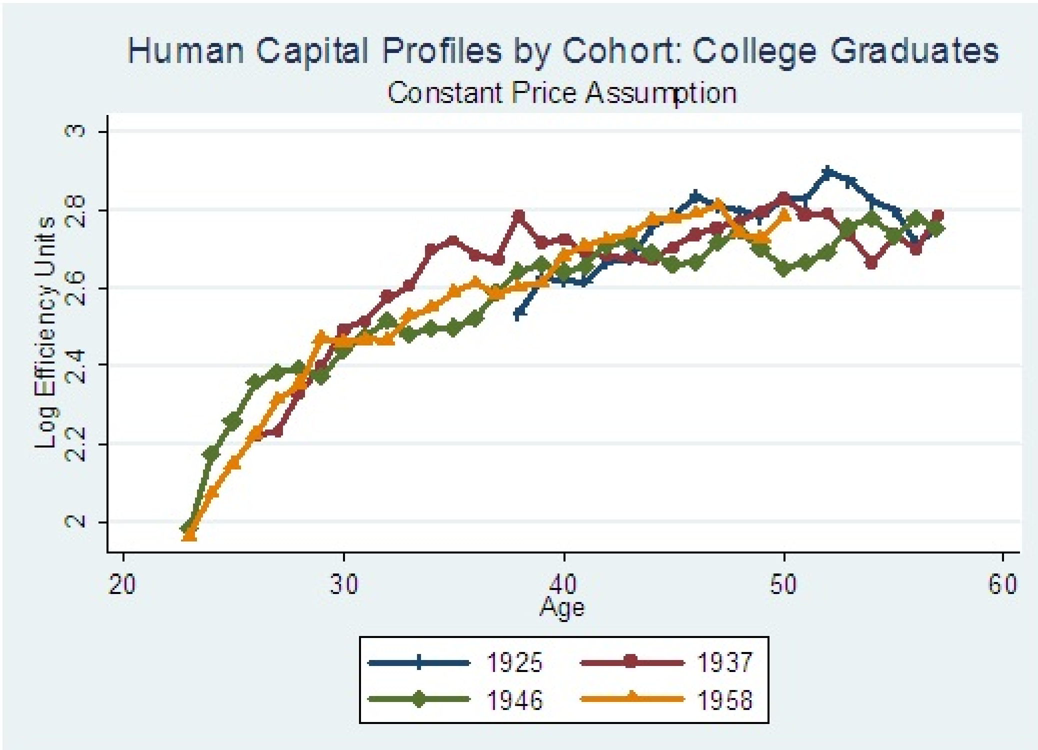

Figure 12b

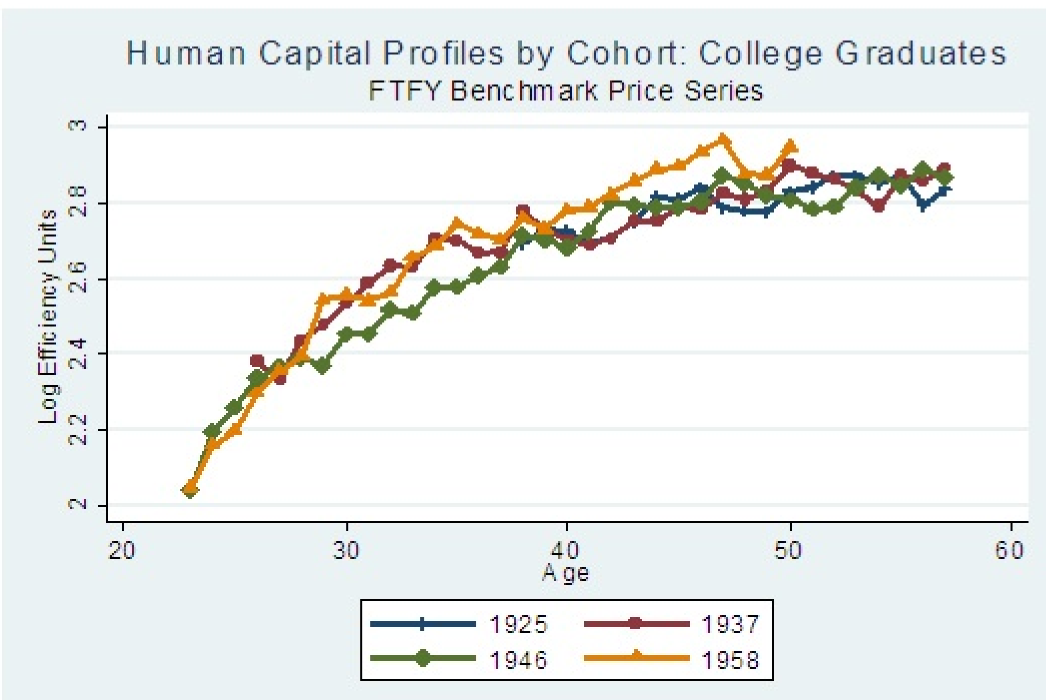

Figure A1a

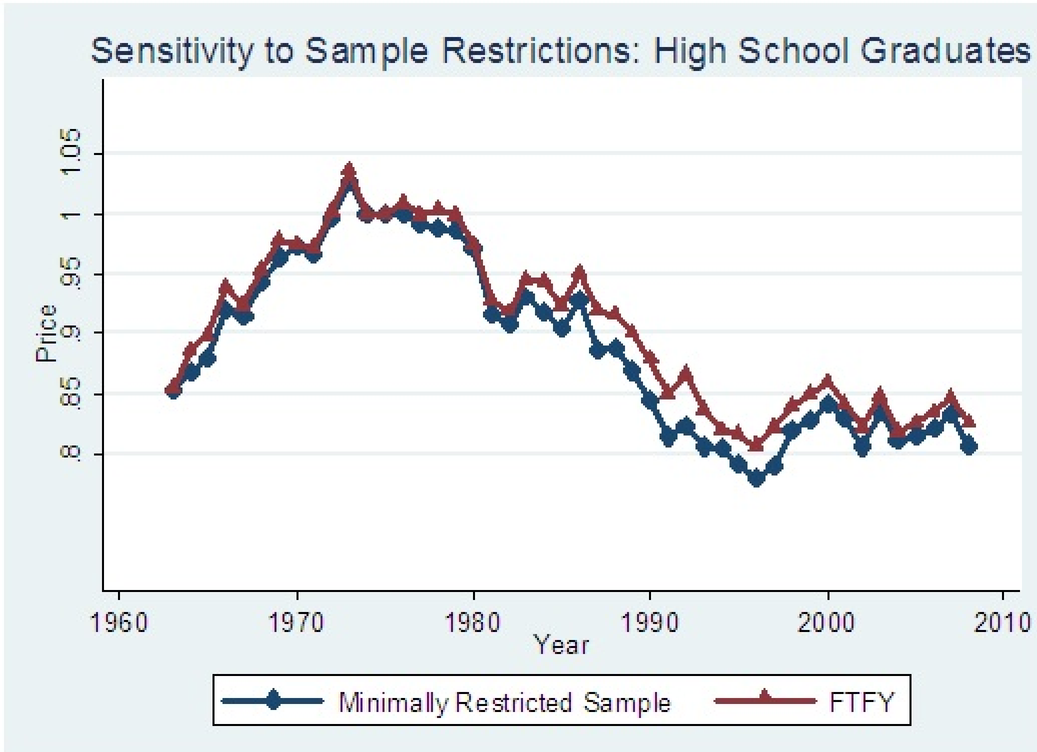

Figure A1b

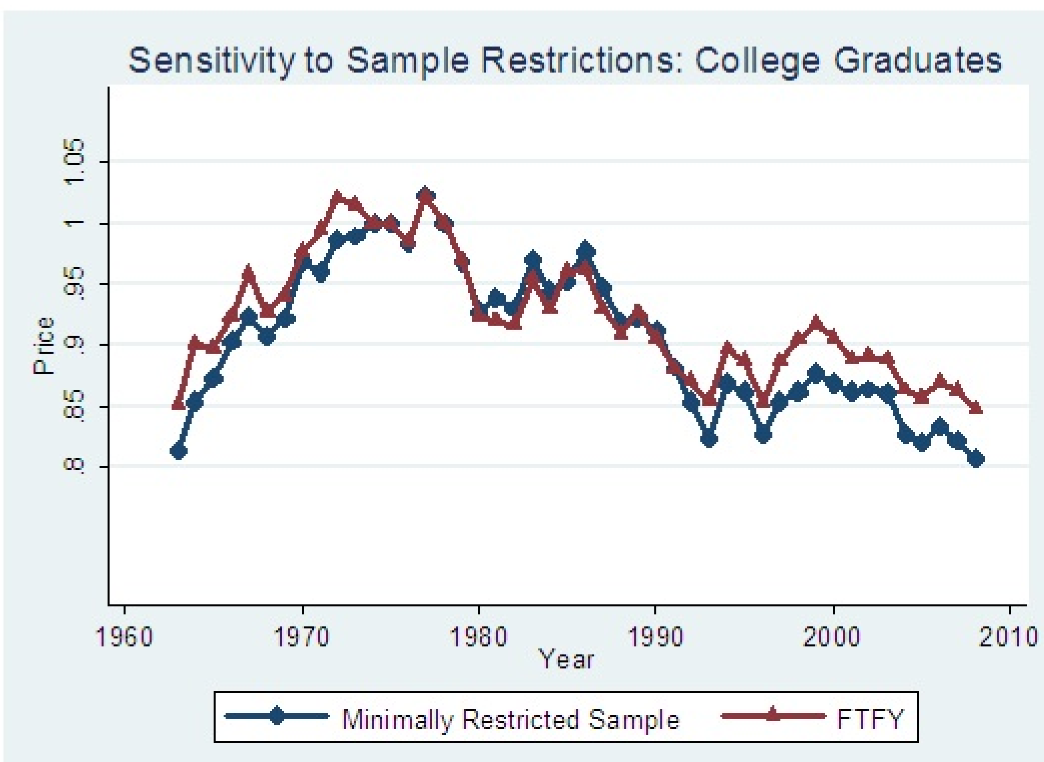

Figure A2a

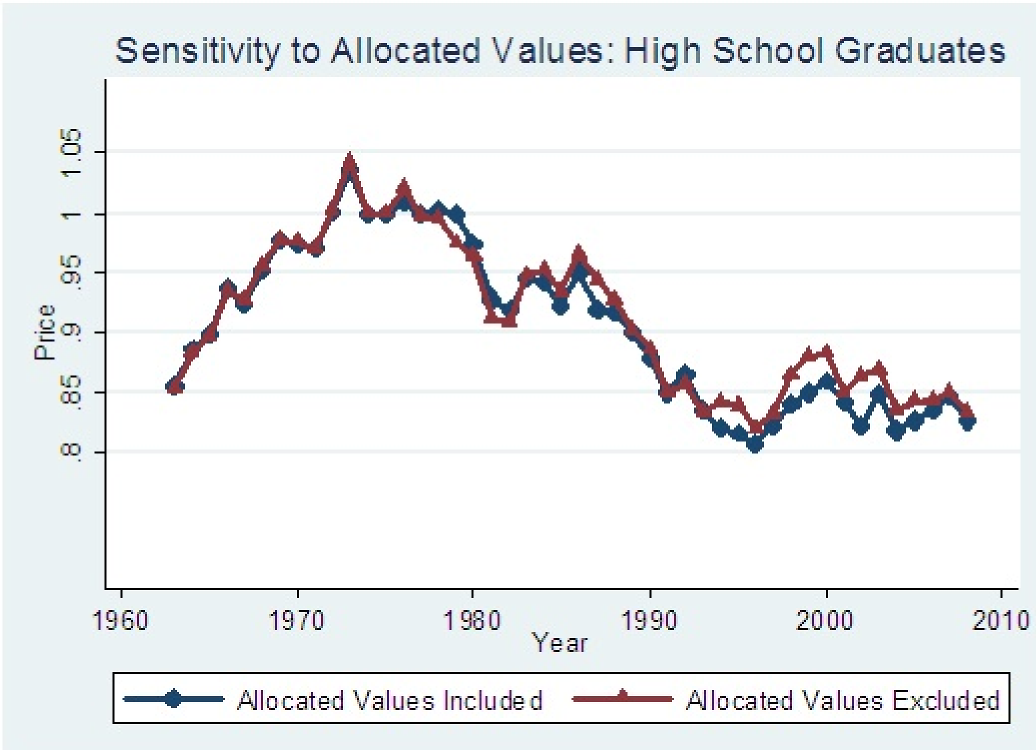

Figure A2b

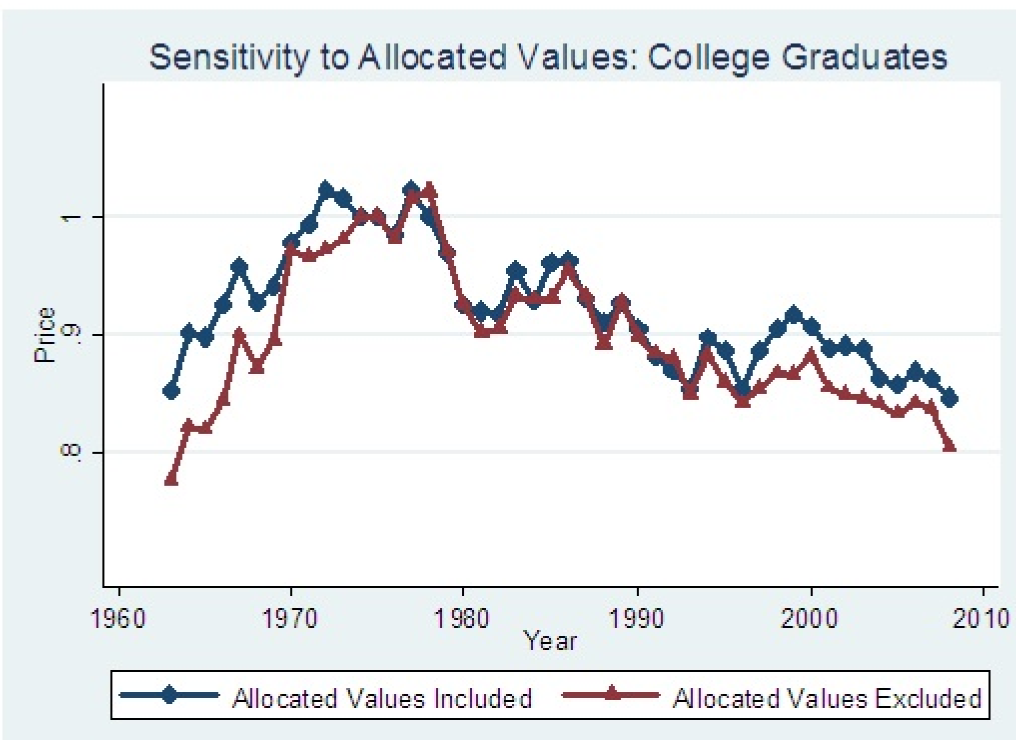

Figure A3

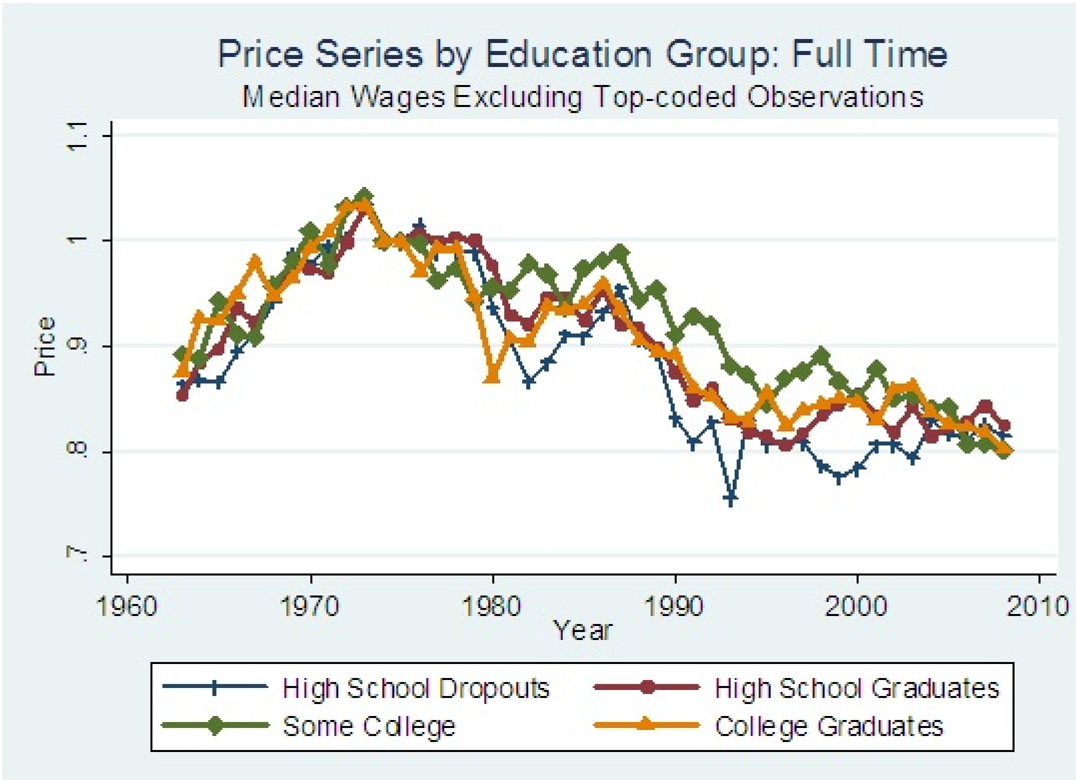

Figure A4a

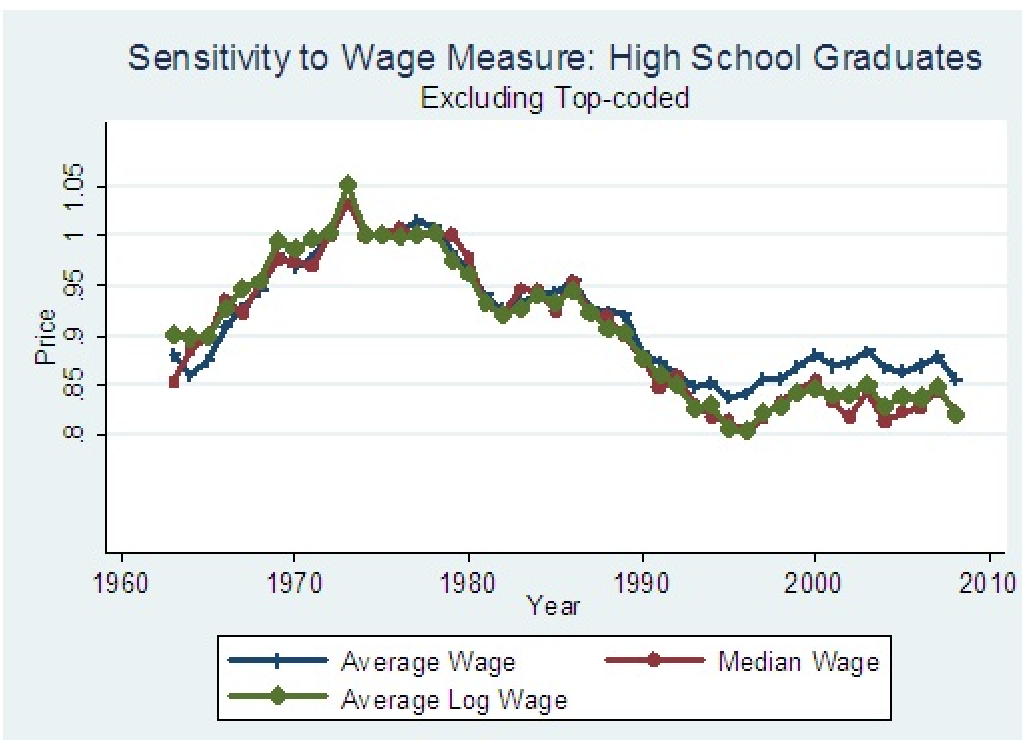

Figure A4b

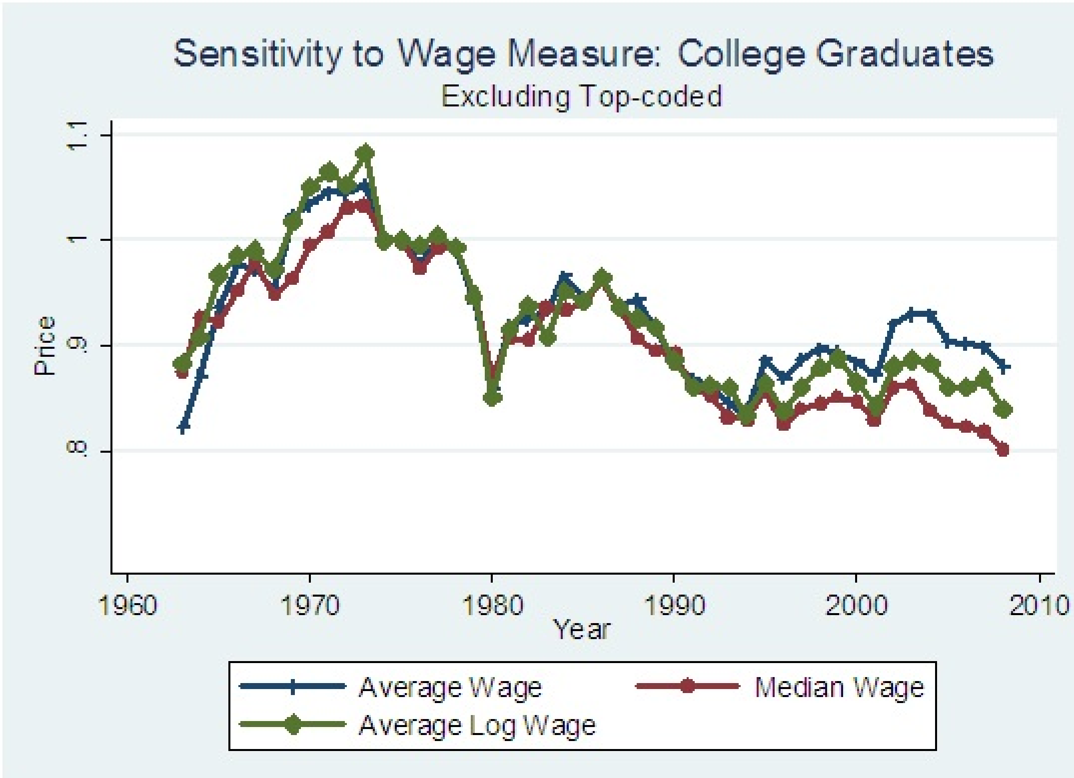

Figure A5