

NBER WORKING PAPER SERIES

A QUANTITATIVE ANALYSIS OF THE EVOLUTION OF THE U.S. WAGE DISTRIBUTION:
1970-2000

Fatih Guvenen
Burhanettin Kuruscu

Working Paper 13095
<http://www.nber.org/papers/w13095>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2007

This paper was prepared for the 2009 NBER Macroeconomics Annual. For helpful discussions, we thank Orazio Attanasio, Richard Blundell, Satyajit Chatterjee, Dean Corbae, Gueorgui Kambourov, Lawrence Katz, Per Krusell, Edward Prescott, Tom Sargent, Robert Shimer, Kjetil Storesletten, Aleh Tsyvinski, Amir Yaron, and especially Daron Acemoglu, David Autor, Francesco Caselli, and Steve Davis, as well as the seminar participants at Carnegie-Mellon, European University Institute, Georgetown, Goethe University Frankfurt, HEC Lausanne, Humboldt, LSE, NYU, Rice, SUNY Stony Brook, University College London, Michigan, Rochester, Toronto, Toulouse, Turkish Central Bank, UT-Austin, Yale, York, University of Zurich, the Federal Reserve Banks of Atlanta and Philadelphia, NBER's Economic Fluctuations and Growth Meeting in Cambridge, the Minnesota Macro Workshop, the Labor Supply and Lifecycle Productivity Workshop organized by the Bank of Canada, the 20th Canadian Macro Study Group Workshop, CEPR's 2007 ESSIM Conference, and the SED conference. Guvenen acknowledges financial support from the National Science Foundation (SES-0649437). All remaining errors are our own. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

© 2007 by Fatih Guvenen and Burhanettin Kuruscu. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

A Quantitative Analysis of the Evolution of the U.S. Wage Distribution: 1970-2000
Fatih Guvenen and Burhanettin Kuruscu
NBER Working Paper No. 13095
May 2007, Revised May 2009
JEL No. E21,E25,J24,J31

ABSTRACT

In this paper, we construct a parsimonious overlapping-generations model of human capital accumulation and study its quantitative implications for the evolution of the U.S. wage distribution from 1970 to 2000. A key feature of the model is that individuals differ in their ability to accumulate human capital, which is the main source of wage inequality in this model. We examine the response of this model to skill-biased technical change (SBTC), which is modeled as an increase in the trend growth rate of the price of human capital starting in the early 1970s. The model displays behavior that is consistent with several important trends observed in the US data, including the rise in overall wage inequality; the fall and subsequent rise in the college premium, as well as the fact that this behavior was most pronounced for younger workers; the rise in within-group inequality; the stagnation in median wage growth; and the small rise in consumption inequality despite the large rise in wage inequality. We consider different scenarios regarding how individuals' expectations evolve during SBTC. Specifically, we study the case where individuals immediately realize the advent of SBTC (perfect foresight), and the case where they initially underestimate the future growth of the price of human capital (pessimistic priors), but learn the truth in a Bayesian fashion over time. Lack of perfect foresight appears to have little effect on the main results of the paper. Overall, the model shows promise for explaining a diverse set of wage distribution trends observed since the 1970s in a unifying human capital framework.

Fatih Guvenen
Department of Economics
University of Minnesota
4-151 Hanson Hall
1925 Fourth Street South
Minneapolis, MN, 55455
and NBER
guvenen@umn.edu

Burhanettin Kuruscu
University of Texas at Austin
Department of Economics
Office BRB 3.122
Austin, Texas 78712
kuruscu@gmail.com

A Quantitative Analysis of the Evolution of the U.S. Wage Distribution: 1970–2000*

Fatih Guvenen[†] Burhanettin Kuruscu[‡]

May 22, 2009

Abstract

In this paper, we construct a parsimonious overlapping-generations model of human capital accumulation and study its quantitative implications for the evolution of the U.S. wage distribution from 1970 to 2000. A key feature of the model is that individuals differ in their ability to accumulate human capital, which is the main source of wage inequality in this model. We examine the response of this model to skill-biased technical change (SBTC), which is modeled as an increase in the trend growth rate of the price of human capital starting in the early 1970s. The model displays behavior that is consistent with several important trends observed in the US data, including the rise in overall wage inequality; the fall and subsequent rise in the college premium, as well as the fact that this behavior was most pronounced for younger workers; the rise in within-group inequality; the stagnation in median wage growth; and the small rise in consumption inequality despite the large rise in wage inequality. We consider different scenarios regarding how individuals' expectations evolve during SBTC. Specifically, we study the case where individuals immediately realize the advent of SBTC (perfect foresight), and the case where they initially underestimate the future growth of the price of human capital (pessimistic priors), but learn the truth in a Bayesian fashion over time. Lack of perfect foresight appears to have little effect on the main results of the paper. Overall, the model shows promise for explaining a diverse set of wage distribution trends observed since the 1970s in a unifying human capital framework.

*This paper was prepared for the 2009 *NBER Macroeconomics Annual*. For helpful discussions, we thank Orazio Attanasio, Richard Blundell, Satyajit Chatterjee, Dean Corbae, Gueorgui Kambourov, Lawrence Katz, Per Krusell, Edward Prescott, Tom Sargent, Robert Shimer, Kjetil Storesletten, Aleh Tsyvinski, Amir Yaron, and especially Daron Acemoglu, David Autor, Francesco Caselli, and Steve Davis, as well as the seminar participants at Carnegie-Mellon, European University Institute, Georgetown, Goethe University Frankfurt, HEC Lausanne, Humboldt, LSE, NYU, Rice, SUNY Stony Brook, University College London, Michigan, Rochester, Toronto, Toulouse, Turkish Central Bank, UT-Austin, Yale, York, University of Zurich, the Federal Reserve Banks of Atlanta and Philadelphia, NBER's Economic Fluctuations and Growth Meeting in Cambridge, the Minnesota Macro Workshop, the Labor Supply and Lifecycle Productivity Workshop organized by the Bank of Canada, the 20th Canadian Macro Study Group Workshop, CEPR's 2007 ESSIM Conference, and the SED conference. Guvenen acknowledges financial support from the National Science Foundation (SES-0649437). All remaining errors are our own.

[†]University of Minnesota, Federal Reserve Bank of Minneapolis, and NBER; guvenen@umn.edu; www.econ.umn.edu/~guvenen.

[‡]University of Texas at Austin; kuruscu@eco.utexas.edu; <http://www.eco.utexas.edu/~kuruscu>.

1 Introduction

The U.S. economy, and the labor market in particular, has experienced substantial changes since the early 1970s. Among the most notable of the trends observed during this period was a rise in overall wage inequality that started slowly in the 1970s, but accelerated substantially in the 1980s. Perhaps surprisingly, however, wage inequality *between* education groups (i.e., the college premium) followed a different pattern: it fell precipitously during the 1970s, but then also rose substantially in most of the subsequent two decades. Despite these big changes in cross-sectional measures of wage dispersion, the rise in consumption inequality (which is a proxy for the inequality in lifetime wage incomes) was very much muted. Finally, in addition to these trends in inequality, macroeconomists have documented a sharp slowdown in aggregate labor productivity growth, and labor economists have found a parallel stagnation in median wage growth, which both started around 1973 and lasted until the mid-1990s.¹

While these trends have typically been documented by economists in different fields, studying different questions, they in fact all point to *changes in the moments of the U.S. wage distribution* during this time. Motivated by decades of research emphasizing the central role of human capital accumulation for the determination of wages, it seems natural to wonder whether the human capital theory can shed light on these developments. Therefore, the goal of this paper is to investigate how much mileage one can get toward explaining these phenomena using a parsimonious overlapping-generations model of human capital accumulation.

The model we construct has the following features. Individuals are born with a fixed endowment of “raw labor” (health, strength, etc.), but are able to accumulate “human capital” (skills, knowledge, etc.) over the life cycle. Raw labor and human capital earn separate wages in the labor market, and each individual supplies both of these factors of production at competitively determined wages. In a given period, an individual is either employed full-time or is enrolled in school (accumulating human capital full-time). However, while employed, an individual can choose to allocate any fraction of his time—subject to an upper bound—to human capital accumulation. We assume that skills are general and the labor market is competitive. As a result, the cost of this on-the-job investment will be completely borne by workers, and firms will adjust the hourly wage rate downward by the fraction of time workers spend learning new skills (Becker (1964)). Thus, the cost of human capital investment is given by these forgone earnings.

An important difference of the framework outlined above from the standard human capital model (e.g., Ben-Porath (1967)) is the introduction of raw labor as a second factor of production. This “two-factor” structure serves a key role in the results of this paper. This is because the standard Ben-Porath model lacks a well-defined notion of “returns-to-skill,” which is essential for studying *skill*-biased technical change as we do in this paper. As we elaborate in Section 2.4, while the Ben-Porath model has a “price of human capital,” in a

¹For extensive documentation of these trends, see Bound and Johnson (1992), Katz and Murphy (1992), Juhn, Murphy, and Pierce (1993), Card and Lemieux (2001), Autor, Katz, and Kearney (2005a,b), Krueger and Perri (2006), and the excellent surveys by Katz and Autor (1999) and Acemoglu (2002).

stationary world, a higher price level (or wage) affects the cost and benefit of investment in exactly the same way, leaving the trade-off, and therefore the investment decision, unaffected.² In contrast, in the present model, the marginal cost of investment (i.e., forgone earnings) is proportional to the prices of both human capital *and* raw labor, whereas the marginal benefit of investment is proportional only to the former. Therefore, a high price of human capital increases the benefit more than the cost, resulting in higher investment.

A second key feature we introduce into this framework is heterogeneity in the ability to accumulate human capital. As a result, individuals differ systematically in the amount of investment they undertake, and consequently, in the growth rate of their wages over the life cycle. Although the idea that individuals may differ in their ability is not new, here we are motivated by recent microeconomic evidence which finds that such heterogeneity is substantial (Baker (1997), Guvenen (2007, 2009), and Huggett, Ventura, and Yaron (2006a,b)). For the parameters chosen, the model generates a large rise in within-cohort wage inequality over the *life cycle*, quantitatively consistent with the values reported in these studies.

The production side of the economy is modeled as an aggregate constant elasticity of substitution (CES) technology that takes raw labor and human capital as inputs. The third key feature in the model, and the driving force behind the non-stationary changes during this period, is skill-biased technical change (SBTC)—modeled here as a rise in the price of human capital relative to raw labor—that occurs starting in the early 1970s.³ In the baseline model, we assume that individuals do not anticipate SBTC before it happens, but have perfect foresight about the future once it starts. (We relax this assumption later.) It is important to point out that the specification of the production function here departs from the existing literature, which typically assumes a CES production function that takes the labor supplied by workers with high and low education as its two inputs (cf., Katz and Murphy (1992), Juhn, Murphy, and Pierce (1993)). Furthermore, because SBTC is typically modeled in these studies as an increase in the relative demand for educated workers, it creates variation in wages *between* education groups, but does not have a differential impact on individuals *within* each group. Instead, in the present framework all individuals supply both factors of production, and therefore, simultaneously lose from the fall in the price of raw labor and gain from the increase in the price of human capital. Moreover, since individuals differ in both age and ability level, these gains and losses are distributed differently across the population (including within each education group), which creates rich dynamics in the evolution of the wage distribution. This feature allows us to study both *between-group* and *within-group* wage inequality in a single framework.

The model is calibrated to match some key moments of the wage distribution before SBTC takes effect in 1970. We then systematically examine the implications of the model for the evolutions after 1970 of (i) the *second* moments (measures of inequality), (ii) the *first*

²To increase investment incentives in the Ben-Porath model, one would need to assume an acceleration in the growth rate of the price of human capital. But in a model with a single factor of production, this would also mean an acceleration in total factor productivity (TFP) growth rates during the period since 1970s, which is clearly counterfactual.

³For empirical evidence on SBTC, see, for example, Berman, Bound, and Griliches (1994), Autor, Katz, and Krueger (1998), Machin and Van Reenen (1998), and the survey by Acemoglu (2002).

moments (measures of average wages), and (iii) *lifetime* wage distribution (proportional to consumption in the model).

First, in the model overall wage inequality (measured by the cross-sectional variance of log wages) rises slowly in the 1970s, but substantially faster in the 1980s, as in the data. The slow rise early on is closely linked to the decline in *between*-group inequality (college premium) during the same time, which brings the wages of different education groups closer to each other, thereby compressing the overall distribution. However, counteracting this force is a rise in *within*-group wage inequality, which prevents overall inequality from falling in the 1970s. Therefore, the model is consistent with the behavior of both between-group and overall wage inequality during the 1970s.

The mechanism behind the non-monotonic behavior of the college premium can be explained as follows. Essentially, because human capital investment is a forward-looking decision, individuals increase their investment immediately after SBTC begins. Moreover, because college graduates have higher learning ability (by self-selection) than those with lower education, their investment increases more strongly, which increases their forgone earnings and therefore reduces their relative wages in the short run. Over time, this higher investment begins to pay off, and the college premium starts to grow rapidly after the initial decline. As can be expected from this discussion, this mechanism is stronger for younger individuals who face a longer horizon and, thus, expect higher benefits from SBTC. Therefore, the college premium falls and rises substantially for young workers (but much less for the old), consistent with this behavior in the U.S. data (cf., Murphy and Welch (1992), and Card and Lemieux (2001)). Finally, the model is also consistent with the behavior of the relative wages of each education-experience group during this period (in particular, with the steepening (flattening) of cross-sectional profiles for high school (college) graduates in the 1980s) documented by Katz and Murphy (1992) and Bound and Johnson (1992).

We then turn to the behavior of labor productivity growth (and median wage growth) in the model, which both fall sharply in the early 1970s in response to SBTC and then recover very slowly. The slowdown is quantitatively large, with the median wage growth rate between 1970 and 1995 averaging about half its value before 1970. Labor productivity slowdown is sharper in the short run, but the recovery is also faster, averaging 80 percent of its value prior to SBTC in the period from 1970 to 1995.

Several authors have documented that consumption inequality in the United States has not increased nearly as much as wage inequality, especially at the upper tail (cf., Autor, Katz, and Kearney (2004), and Krueger and Perri (2006)). Turning to the model, inequality in (the present discounted value of) lifetime income barely rises after 1970. The intuition for this result can be understood by noting that wage inequality in the model increases due to the fanning out of lifecycle wage profiles (see figure 12). As a result, those individuals whose wages rise the most later in life are exactly those whose wages remain lower early in life due to increased human capital investment, which then keeps the change in lifetime incomes small. This mechanism is consistent with the observation that individuals in many high-skill occupations (doctors, engineers, professors, etc.), whose incomes have risen tremendously in the last decades, have had to go through longer periods of education, training, internship,

certification, and so on to be able to attain those high income levels.

The amount of average on-the-job investment implied by the model is 2.9 hours a week (assuming a 40-hour workweek) in 1970 and rises to 5.1 hours a week in 1995. While ideally, one would also want to compare this and other implications for human capital accumulation with the data, this approach is not straightforward. This is because, as is well understood since Becker (1964), Ben-Porath (1967), and others, on-the-job investment in human capital theory is a much broader concept than the relatively limited notion of training directly provided by firms, which makes the total amount of such investment notoriously difficult to measure, let alone quantifying how much it has *changed* over time for different groups in the population. For example, Barron, Berger, and Black (1997) piece together information from six different data sources and estimate that total investment on the job is at least seven times larger than the formal training provided by firms. The alternative approach of focusing on educational attainment does not seem to be very fruitful either. This is because, although formal education and on-the-job investment are likely to be positively related on average, this is not always the case, especially in response to major technological changes such as SBTC. For example, we provide an extension of the present framework in Guvenen and Kuruscu (2007, section 4), which retains many of the plausible implications for the wage distribution, but implies that college enrollment *falls*—whereas on-the-job investment *rises*—in the short run after SBTC. That extension, however, makes the model substantially more complicated and is therefore beyond the scope of this paper. Thus, we keep the focus of this paper on the evolution of the wage distribution and do not venture into providing a model to study the details of educational attainment. See section 4.4 for further discussion of these points.

As noted above, in the baseline model we assume that individuals have perfect foresight after SBTC begins. While this assumption is common in the literature, there are good reasons to question the robustness of the results to such a stark assumption. Therefore, in sections 3 and 4, we also study the implications of the model when individuals have imperfect foresight and form initial beliefs about the future growth rate of factor prices (SBTC), which they then update over time in a Bayesian fashion. We consider the case where these initial beliefs are unbiased as well as when they are very pessimistic (such that the average individual in 1970 forecasts no SBTC in the future, and half of the population in fact expects the price of human capital to continually fall in the future). Although some new and interesting channels become operational with imperfect foresight, for plausible parameter values these effects turn out to be too small and do not alter the conclusions drawn in this paper. Furthermore, in section 4.1.3 we also show that in certain cases, imperfect foresight can *strengthen* the mechanisms outlined for the deterministic model above.

In a companion paper, Guvenen and Kuruscu (2007), we theoretically study a slightly simplified version of the baseline model in this paper. Whereas in that paper we establish several results about the behavior of the wage distribution during SBTC theoretically, the main contribution of the present paper is to provide a comprehensive assessment of the model's ability to explain the observed trends quantitatively.

A precursor to our paper is Heckman, Lochner, and Taber (1998). To our knowledge, their paper was the first to emphasize that with human capital accumulation, skill prices

and observed wages differ, and that this could be important for understanding the recent rise in wage inequality. This observation also plays an important role in our model. However, the present paper also differs from theirs in several important respects. Two of these have already been mentioned above: *One*, the present paper extends the Ben-Porath model to allow for returns-to-skill; and *two*, a central thesis here is that individuals differ significantly in their ability to accumulate human capital, which is not the case in that paper. As a result, for example, the college premium falls significantly in our model even when the production function is linear, whereas imperfect substitution in the production function and changes in cohort sizes over time are essential for this result in that paper. In addition, in the *long run* after SBTC all measures of inequality *increase* in the present paper, whereas many of them—overall wage inequality, college-premium, within-group inequality—*fall* in that paper. Finally, we also examine the robustness of our results to the lack of perfect foresight. We discuss a number of other differences between the two models later in the paper.

This paper is also related to the macroeconomics literature that studies the slowdown in labor productivity. Important examples include Hornstein and Krusell (1996), Greenwood and Yorukoglu (1997), Caselli (1999), and Violante (2002). Greenwood and Yorukoglu (1999) emphasize the role of skill in facilitating the adoption of new technologies. They argue that the advent of computer technologies in the 1970s presented such a change, which increased the wages of skilled workers and resulted in a productivity slowdown due to the time it takes to utilize the new technologies effectively. Hornstein and Krusell (1996) also make a similar observation, but add that the acceleration in quality improvements during this period has exacerbated measurement problems, further reducing measured productivity growth. Caselli (1999) studies a model where differences in innate ability and newer technologies that are more costly to learn than existing ones result in a rising skill premium. Violante (2002) develops a model of within-group inequality, in which vintage-specific skills, embodied technological acceleration, and labor market frictions combine to generate rising inequality. These papers share the feature that technical change is embodied in new machines; in the present paper, however, it is disembodied. It is possible to argue that both types of technological changes have been taking place during this period, so the mechanisms emphasized in these papers are complementary to ours.

The paper is organized as follows. The next section presents the model, and section 3 discusses the calibration. Section 4 contains the results of the baseline model with perfect foresight as well as the versions with imperfect foresight. Section 5 concludes.

2 Baseline Model

The economy consists of overlapping generations of individuals who live for S years. There is no population growth and we normalize the population size to 1, implying that the measure of s -year old individuals, $\mu(s)$, is equal to $1/S$.

2.1 A Two-Factor Model of Human Capital Accumulation

Individuals begin life with an endowment of “raw labor” (i.e., strength, health, etc.) which is constant over the life cycle, and are able to accumulate “human capital” (skills, knowledge, etc.) over the life cycle, which is the only skill that can be accumulated in this economy. There is a continuum of individuals in every cohort, indexed by $j \in [0, 1]$, who differ in their ability to accumulate human capital, denoted by A^j (also referred to as their “type”). Although, in some cases below, we will allow individuals to also differ in their raw labor endowment as well as in their beliefs, the heterogeneity in ability will be the crucial source of heterogeneity in the model.

Each individual has one unit of time endowment in each period that can be allocated between producing output and accumulating human capital. Let l denote raw labor and $h_{s,t}^j$ denote the human capital in period t of an s -year-old individual of type j . We assume that raw labor and human capital earn separate wages in the labor market, and each individual supplies both of these factors of production at competitively determined wage rates, denoted by $P_{L,t}$ and $P_{H,t}$, respectively.⁴

Individuals accumulate human capital according to the following technology:

$$h_{s+1,t+1}^j = h_{s,t}^j + \underbrace{A^j((\theta_{L,t}l + \theta_{H,t}h_{s,t}^j)i_{s,t}^j)^\alpha}_{Q_{s,t}^j}, \quad (1)$$

where $i_{s,t}^j$ is the fraction of time devoted to human capital investment, henceforth referred to as “investment time”; and $Q_{s,t}^j$ is the newly produced human capital, which will be referred to simply as “investment” in the rest of the paper. According to this formulation, new human capital is produced by combining the existing stocks of raw labor and human capital with the available investment time.⁵

A key parameter in this specification is A^j , which determines the productivity of learning. Due to the heterogeneity in A^j , individuals will differ systematically in the amount of investment they undertake, and consequently, in the growth rate of their wages over the life cycle. Another important parameter is $\alpha \in [0, 1]$, which determines the degree of diminishing marginal returns in the human capital production function. A low value of α implies higher diminishing returns, in which case it is optimal to spread investment over time. In contrast, when α is high the marginal return on investment does not fall quickly, and investment becomes bunched over time. In the extreme case when $\alpha = 1$, individuals either spend all their time on investment ($i_{s,t}^j = 1$) or none at all in a given period.

⁴The structure we have in mind is not one where an individual works at manual tasks (using raw labor only) some fraction of the time and at cognitive (or skill-intensive) tasks at other times. Instead, the worker employs both factors of production simultaneously in producing output. For example, a college professor uses both his/her body and knowledge/skills at the same time when teaching, although probably at different proportions than a farmer, an auto mechanic, or a brain surgeon.

⁵The dependence of the weights in the human capital production function on t is to stress that these could be time-varying.

Investment in human capital takes place on the job as long as it does not exceed a fraction χ of an individual's time endowment in a period. If the individual wants to invest more, he enrolls in college and invests 100 percent of his time. Thus, the choice set for investment time is $i_{s,t}^j \in [0, \chi] \cup \{1\}$, which is non-convex when $\chi < 1$. An upper bound less than 100 percent on on-the-job investment seems plausible as it could arise, for example, if the firm incurs fixed costs for employing each worker (administrative burden, cost of office space, etc.), or due to minimum wage laws.⁶

We assume that skills are general (i.e., not firm-specific) and labor markets are competitive. As a result, the cost of human capital investment is completely borne by workers, and firms adjust the hourly wage rate downward by the fraction of time invested on the job (Becker 1964). Then, the observed wage income of an individual is given by

$$w_{s,t}^j = [P_{L,t}l + P_{H,t}h_{s,t}^j] (1 - i_{s,t}^j) = \underbrace{[P_{L,t}l + P_{H,t}h_{s,t}^j]}_{\text{Potential earnings}} - \underbrace{[P_{L,t}l + P_{H,t}h_{s,t}^j] i_{s,t}^j}_{\text{Cost of investment}}, \quad (2)$$

where $[P_{L,t}l + P_{H,t}h_{s,t}^j]$ is the “potential earnings” of an individual—that is, the income an individual would earn if he spent all his time producing for his employer. Therefore, wage income can be written as the potential earnings minus the “cost of investment,” which is simply the forgone earnings while individuals are learning new skills. Since labor supply is inelastic, $w_{s,t}^j$ is also just a scaled version of the individual's observed “wage rate.”

2.2 A CES Aggregate Production Function

Let the aggregate factors used in production at a point in time be defined as

$$L_t^{net} \equiv \sum_{s=1}^S \mu(s) \int_j l (1 - i_{s,t}^j) dj, \quad \text{and} \quad H_t^{net} \equiv \sum_{s=1}^S \mu(s) \int_j h_{s,t}^j (1 - i_{s,t}^j) dj, \quad (3)$$

where the integrals are thus taken over the distribution of individuals of all types and ages. The superscripts in H_t^{net} and L_t^{net} stress that these variables measure the actual amounts of each factor used in production, that is, *net* of the time allocated to human capital investment. The aggregate firm uses these two inputs to produce a single good, denoted by Y , according to the familiar CES production function:

$$Y = Z ([\theta_L L^{net}]^\rho + [\theta_H H^{net}]^\rho)^{1/\rho}, \quad (4)$$

⁶In addition to its plausibility, such an upper bound is also important for a meaningful quantitative analysis. Otherwise, with a continuum of ability levels, there will be some individuals who invest slightly less than 100 percent of their time, appearing as employed while earning a wage income very close to zero. Because many of the statistics we analyze below involve the logarithm of wage rates as well as the variances of these logarithms, even a small number of such individuals can easily wreak havoc with the quantitative exercise.

where $\rho \leq 1$, and Z is the total factor productivity (TFP). For simplicity we assume that capital is not used in production. Notice that human capital and raw labor enter the aggregate production function and human capital function with the same weights (compare equations (1) and (4)).⁷

The firm solves a static problem by hiring factors from households every period to maximize its profit: $Y - P_L L^{net} - P_H H^{net}$. Using the optimality conditions for this problem, the price of human capital *relative* to raw labor can be expressed as

$$\frac{P_H}{P_L} = \left(\frac{\theta_H}{\theta_L} \right)^\rho \left(\frac{H^{net}}{L^{net}} \right)^{\rho-1}. \quad (5)$$

While the aggregate production function has the same CES form commonly used in the literature, its inputs are different from what is typically assumed. In most previous work, H^{net} and L^{net} denote the number of workers with college and high school education, respectively. Therefore, a change in the price of H^{net} relative to L^{net} has the *same effect* on all individuals *within* an education group. As a result, the college premium is simply equal to P_H/P_L and satisfies the relationship in (5). A key implication of this equation is that a rise in the relative supply of high-skill workers will reduce the college premium. Several authors have emphasized this link to argue that the fall in the college premium during the 1970s resulted from the rapid increase in the supply of college-educated workers (cf., Katz and Murphy (1992), Juhn, Murphy, and Pierce (1993), and Heckman et al. (1998)). In contrast, in the present model, all workers have some endowment of human capital (which varies by ability and age) and l (which is the same for all), and every worker contributes to both factors of production. Therefore, a change in the price of H^{net} relative to L^{net} affects all individuals differently depending on their ability level as well as their age. As a result, the present model allows us to study both within-group and between-group wage inequality in a unifying framework. Moreover, as we show below, the college premium is now very different from P_H/P_L .

An important special case arises when $\rho = 1$. In this case, human capital and raw labor become perfectly substitutable, and the relative wage in equation (5) reduces to $P_H/P_L = \theta_H/\theta_L$. Therefore, this assumption eliminates the link between the relative supply of high-skill labor and the college premium mentioned above. To highlight the role of the mechanism proposed in this paper for the college premium, we make this assumption from this point on. Furthermore, in the appendix we provide evidence indicating that this is an empirically plausible choice for the curvature for this type of production function.⁸

⁷We have experimented with two other specifications of the human capital accumulation function that also seemed a priori plausible. In particular, we write the human capital function as $A^j((\lambda_{L,t}l + \lambda_{H,t}h_s^j)v_s^j)^\alpha$ and consider (i) weights that remain constant through SBTC: $\lambda_{L,t} = \bar{\lambda}_L$ and $\lambda_{H,t} = \bar{\lambda}_H$; and (ii) no role for raw labor in human capital production: $\lambda_{L,t} = 0$ and $\lambda_{H,t} = 1$. The first case had implications qualitatively similar to the baseline model described here, while the second displayed some implausible behavior even in steady state (that is, without SBTC). Overall, the simplest specification used here (originally proposed and studied in Guvenen and Kuruscu (2007)) also turned out to have the most plausible quantitative implications.

⁸In the NBER working paper version of this article (No. 13095, Appendix B), we also analyze the more

2.3 SBTC and Bayesian Learning about Skill Prices

Skill-*neutral* technological progress takes place at a constant rate: $Z_{t+1} = (1 + g) Z_t$. The focus of this paper is on skill-*biased* changes in technology, which we model as follows. The productivity of each factor follows a random walk with drift in levels:⁹

$$\begin{aligned}\theta_{H,t+1} &= \kappa + \theta_{H,t} + \varepsilon_{t+1}, & \text{and} \\ \theta_{L,t+1} &= -\kappa + \theta_{L,t} - \varepsilon_{t+1},\end{aligned}\tag{6}$$

where $\varepsilon \sim N(0, \sigma_\varepsilon^2)$. There are several points to note in this specification. First, the innovations to the productivity of each factor sum to zero, so these random shocks only affect the productivity of each factor *relative* to the other. Second, the secular growth rates of the productivity of each factor also sum to zero. Therefore, κ captures the trend growth in skill-biased technical change (SBTC) when it is positive. Putting the two together shows that κ and ε_{t+1} only affect the relative productivity of factors: $\theta_{H,t+1} + \theta_{L,t+1} = \theta_{H,t} + \theta_{L,t}$.

The growth rate of each factor’s productivity is zero (i.e., $\kappa \equiv 0$) up to time 0. Skill-biased technical change is modeled as an unanticipated regime change in the growth rate of human capital’s productivity relative to raw labor. Specifically, $\kappa = \kappa^* > 0$, for $t = 1, \dots, T$; and κ reverts back to zero again for $t > T$. Therefore, SBTC is assumed to last for a finite period of time, which will be motivated below by empirical evidence.

Modeling Expectations of Future. A difficult question that researchers face when studying periods of transition concerns the modeling of individuals’ expectations of the future: what do individuals know and when do they know about it? For example, did individuals realize the advent of skill-biased technical change (or, call it the IT revolution, the information age, etc.) soon after it started in the 1970s? Or were they completely unaware that many companies were rapidly investing in new IT technologies¹⁰ and that the demand for cognitive skills was rising during that time? Perhaps, as seems plausible, some workers were initially more aware than others—because of differences in education level, social networks, occupations, and so on—but over time, all workers have learned about the new regime. But, if so, how fast was this learning process?

To circumvent these issues, the bulk of the previous literature assumes either perfect foresight or myopic expectations about the future. While both of these extreme cases have the appeal of providing simple benchmarks, the questions raised above cannot be easily

general case with $\rho < 1$ and show that it does not alter any of the substantive conclusions reached in the present paper.

⁹An important advantage of the random walk specification in the levels of $\theta_{H,t}$ and $\theta_{L,t}$ (instead of their logarithms) is that it makes the model tractable and allows a closed-form solution for the optimal investment choice in the presence of uncertainty *and* Bayesian learning about future skill prices. Although a potential drawback of this structure is that the productivity of a factor may become negative, for our choices of parameter values, the probability of this outcome will be negligible.

¹⁰See, for example, Greenwood (1997) and Greenwood and Jovanovic (1999) for detailed descriptive evidence on the rapid pace of technology upgrading by firms during this period.

dismissed. Therefore, in the quantitative analysis we will consider a number of alternative scenarios regarding individuals’ beliefs about the future after SBTC. In particular, we allow individuals to differ in their initial beliefs about the evolution of future skill prices—and as an extreme case, we also allow individuals to systematically underestimate the growth of future skill prices (i.e., have pessimistic prior beliefs)—but learn the truth over time in a Bayesian fashion. We now describe these cases in turn.

In the first benchmark, we assume that (i) skill prices evolve in a deterministic fashion (i.e., $\sigma_\varepsilon \equiv 0$), and (ii) as soon as SBTC begins, individuals learn κ^* and T and, therefore, have perfect foresight about the future paths of $\theta_{H,t}$ and $\theta_{L,t}$. We refer to this case as the deterministic benchmark.

In the second benchmark, we allow for imperfect foresight. Specifically, we assume that before time 0, individuals’ forecasts of κ had converged to the true value ($\kappa \equiv 0$) after observing a sufficiently long history of skill prices. Let n denote an index that *uniquely* identifies each individual in the history of this economy (which can be constructed by taking the interaction of ability type j and year of birth t_0). At the time of the shock, each individual (who is in the labor market at that time) receives an initial (private) signal about the future growth rate: $\kappa_0^n = \kappa^* + \eta_0^n$, where $\eta_0^n \sim N(m_\eta, \sigma_\varepsilon^2/v)$ and where n is an index that uniquely identifies each individual in the economy defined above. Therefore, individual n ’s initial forecast is κ_0^n with a precision of v/σ_ε^2 .¹¹ The population average of initial forecasts is $\kappa^* + m_\eta$. Therefore, prior beliefs are unbiased when $m_\eta = 0$ and are “pessimistic” when $m_\eta < 0$. In subsequent periods, individuals observe their own wage realization and update their forecast of κ . Since the source of uncertainty is aggregate, all individuals observe the same path of prices, and their beliefs eventually converge to each other—and hence to the truth. (The details of the Bayesian learning problem and exact formulas are provided in Appendix 6.)

2.4 Individuals’ Lifetime Income Maximization Problem

We first describe individuals’ maximization problem in the deterministic baseline case and then explain how this changes with stochastic skill prices and Bayesian learning.

Deterministic Benchmark. Individuals are able to borrow and lend at a constant interest rate, r , which is sufficient for markets to be complete in this deterministic model. Under complete markets, the consumption-savings decision can be disentangled from the lifetime income maximization problem. Therefore, to study the determination of wages, restricting our attention to the latter problem is sufficient.

Let $V_s^j(h_{s,t}^j, \theta_{H,t})$ denote the present discounted value of lifetime income of individual j who is s years old at time t . The income maximization problem of the agent can be written

¹¹Similarly, the prior beliefs for individuals who enter the economy at $t_0 > 0$ are drawn from a distribution such that they have the same average forecast ($\bar{\kappa}_{t_0}$) and the same precision ($1/\sigma_{\kappa,t_0}^2$) as individuals already in existence in that year. See Appendix 6 for more details.

recursively as

$$V_s^j(h_{s,t}^j; \theta_{H,t}) = \max_{i_{s,t}^j \in [0, \chi] \cup \{1\}} \left[Z_t(\theta_{L,t}l + \theta_{H,t}h_{s,t}^j)(1 - i_{s,t}^j) + \frac{1}{1+r} V_{s+1}^j(h_{s+1,t+1}^j; \theta_{H,t+1}) \right] \quad (7)$$

subject to

$$h_{s+1,t+1}^j = h_{s,t}^j + A^j((\theta_{L,t}l + \theta_{H,t}h_{s,t}^j)i_{s,t}^j)^\alpha, \quad h_0 = 0, \quad (8)$$

$$V_{S+1}^j \equiv 0,$$

where the evolution of factor prices is given by equation (6) and $\sigma_\varepsilon \equiv 0$. We should stress again that our focus on the lifetime income maximization problem does not require the assumption of risk neutrality; any concave utility function implies the same human capital investment behavior in this environment.

Optimal Investment Decision. It is instructive to discuss the qualitative features of investment behavior in this model, which will play important roles in the quantitative results below.

Using equation (1), the opportunity cost of investing an amount $Q_{s,t}^j$ can be written as

$$C^j(Q_{s,t}^j) \equiv (\theta_{L,t}l + \theta_{H,t}h_{s,t}^j)i_{s,t}^j = \left(\frac{Q_{s,t}^j}{A^j} \right)^{1/\alpha}. \quad (9)$$

When there is an interior solution for investment time (specifically, assuming $\chi = 1$ and $i_{s,t}^j < 1$), optimal investment choice satisfies the following first-order condition:

$$C^j(Q_{s,t}^j)' = (\delta\theta_{H,t+1} + \delta^2\theta_{H,t+2} + \dots + \delta^{S-s-1}\theta_{H,t+S-s-1}), \quad (10)$$

where $\delta \equiv (1+g)/(1+r)$ (see Guvenen and Kuruscu (2007) for derivation). The left-hand side of this equation is the marginal cost, and the right-hand side is the marginal benefit (*MB*) of investing in human capital. The latter is the present discounted value of the future stream of wages that is earned by an additional unit of human capital. An important implication of (10) is that an increase in *future* skill prices will immediately affect *current* investment behavior because of the forward-looking nature of this equation.

Remark. A key observation is that optimal investment, $Q_{s,t}^j$, only depends on the level of θ_H —not on the levels of θ_L or Z . This is because the opportunity cost of investment depends on the prices of both raw labor and human capital (see equation (9)), whereas the marginal benefit is only proportional to the price of human capital. As a result, a higher level of θ_H (for example, due to SBTC) increases the marginal benefit more than the marginal cost, resulting in higher investment. This feature is an important difference between the

current framework and the standard Ben-Porath model. In the latter, a higher price of human capital (which is the only factor of production, since there is no raw labor) affects the cost and benefit of investment exactly the same way, leaving the trade-off—and therefore the investment decision—unaffected. It is precisely for this reason that it is difficult to think of the concept of returns-to-skill in that framework, because a higher price of human capital has no effect on the decision to invest. Instead, in the present model, θ_H/θ_L is a measure of returns-to-skill and affects investment in human capital without necessarily implying anything about aggregate productivity (which is captured by Z and also has no effect on investment incentives for the same reason discussed for the Ben-Porath model).¹²

We now proceed to characterize the optimal investment choice, which can be solved for explicitly:

$$Q_{s,t}^j = (A^j)^{1/(1-\alpha)} [\alpha MB]^{\alpha/(1-\alpha)}. \quad (11)$$

This expression highlights the main sources of heterogeneity in this model: (i) individuals with higher learning ability invest more in human capital: $\partial Q_{s,t}^j / \partial A^j > 0$; (ii) more importantly, their investment responds more strongly to SBTC: $\partial^2 Q_{s,t}^j / \partial \theta_{H,t+k} \partial A^j > 0$ (for all $k > 0$); (iii) investment goes down over the life cycle: $\partial Q_{s,t}^j / \partial s < 0$; and finally, younger individuals respond more strongly to SBTC: $\partial^2 Q_{s,t}^j / \partial \theta_{H,t+k} \partial s < 0$.

Price, Investment, and Quantity Effects. Armed with this characterization of optimal investment behavior, we are now ready to discuss how the wage of a typical worker changes in response to SBTC. For $t > 0$, an individual’s wage can be written as

$$w_{s,t}^j = Z_t [(\theta_{L,0} - t\kappa^*)l + (\theta_{H,0} + t\kappa^*)h_{s,t}^j - C^j(Q_{s,t}^j)].$$

The effects of SBTC on wages work through three separate channels, which can be seen from this expression. *First*, individuals will increase their investment immediately in response to SBTC, resulting in a rise in the cost of investment, which will then reduce the wage rate. We call this the “investment effect.” The discussion above makes clear that this effect is strongest (most negative) for those with high ability and/or those who are young.

Second, even in the absence of any response in the individuals’ investment rate (that is, keeping $Q_{s,t}^j$ and $h_{s,t}^j$ fixed), the wage rate will change due to the fact that $\theta_{L,t}$ is falling and $\theta_{H,t}$ is rising over time. We call this the “price effect.” The price effect will be positive (negative) for individuals who have a high (low) stock of human capital relative to their raw labor (that is, those with high (low) ability and/or labor market experience). Moreover, the price effect strengthens over time as $\theta_{H,t}/\theta_{L,t}$ rises during SBTC.

Third, and finally, the increased investment gradually raises the stock of human capital, $h_{s,t}^j$, which in turn gradually raises the wage rate. We call this the “quantity effect.” Notice that the quantity effect is stronger for younger (and/or high-ability) individuals whose

¹²Although an increase in the *growth rate* of Z will increase investment rates, there is no evidence of increased TFP growth rate after the 1970s; in fact, there is ample evidence to the contrary.

investment responds more strongly to SBTC than it is for older (and/or low-ability) individuals. These three effects are crucial for understanding the changes in the wage structure after SBTC, and we will refer to them throughout the paper.

Individuals' Problem with Bayesian Learning. With stochastic skill prices and imperfect foresight, individuals solve

$$V_s^j(h_{s,t}^n; \widehat{\kappa}_t^n, \theta_{H,t}) = \max_{i_{s,t}^n} \left[Z_t(\theta_{L,t}l + \theta_{H,t}h_{s,t}^n)(1 - i_{s,t}^n) + \frac{1}{1+r} E_t(V_{s+1}^j(h_{s+1,t+1}^n; \widehat{\kappa}_{t+1}^n, \theta_{H,t+1})) \right]$$

subject to $V_{S+1}^j \equiv 0$, and equations (6), (8), and (15). The conditional expectation at time t is taken with respect to the information set containing $(\widehat{\kappa}_t^n, \theta_{H,t})$. Notice that choice and state variables are indexed by individual type n (and not simply by ability type j). This is because, even conditional on ability, individuals who observe different initial signals about the path of skill prices will have different investment and human capital levels. The optimality condition for investment choice at time t is

$$C^j(Q_{s,t}^n)' = E(\delta\theta_{H,t+1} + \delta^2\theta_{H,t+2} + \dots + \delta^{S-s-1}\theta_{H,t+S-s-1} | \widehat{\kappa}_t^n, \theta_{H,t}), \quad (12)$$

which only differs from (10) in the appearance here of the expectations operator, which is individual-specific reflecting the different prior beliefs held by different individuals. Substituting the future prices of human capital from equation (6) and some straightforward manipulation yields

$$C^j(Q_{s,t}^n)' = b_{1,s}\theta_{H,t} + b_{2,s}\widehat{\kappa}_t^n, \quad (13)$$

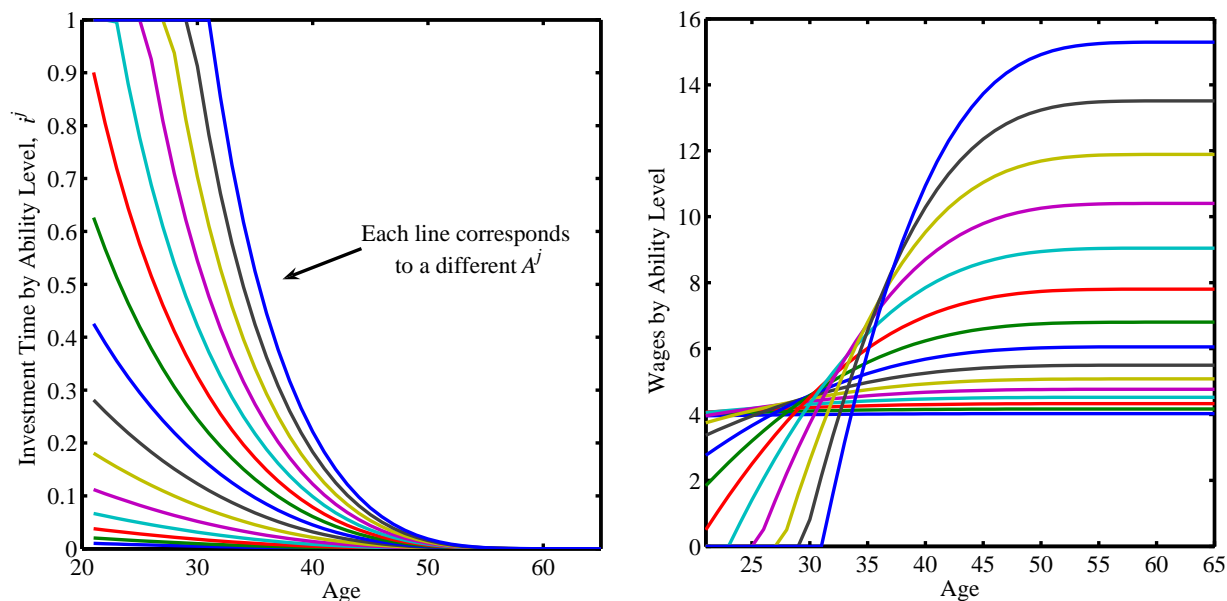
where $b_{1,s}$ and $b_{2,s}$ are some age-dependent positive constants.¹³ The key point to observe here is that optimal investment only depends on the individual's forecast, $\widehat{\kappa}_t^n$, but not on its variance, $\sigma_{\widehat{\kappa}_t^n}^2$. In other words, the *uncertainty* that an individual faces about future prices of human capital plays no role in the optimal investment decision. However, investment behavior in this model will still be different from that in the deterministic benchmark because (i) different individuals have different forecasts $\widehat{\kappa}_t^n$, and (ii) even if $\widehat{\kappa}_t^n \equiv \kappa$ for all individuals in all periods, now $\theta_{H,t}$ is stochastic in equation (13), which will affect the average investment rate in the economy. We return to this point in the quantitative analysis.

2.5 Heterogeneity in Investment and Wages over the Life Cycle

Using the closed-form expressions derived above, we now construct the life cycle profiles of human capital investment and the implied wage paths in steady state (figure 1). The two panels in this figure are generated using the same parameter values as in the calibrated economy of the next section (except that we set $\chi = 1$ here) to highlight the aspects of the life cycle investment behavior that will play a quantitatively important role in our results. The first point to note is the substantial cross-sectional heterogeneity in investment rates observed

¹³ $b_{1,s} = \frac{\delta(1-\delta^{S-s-1})}{1-\delta}$ and $b_{2,s} = \sum_{m=1}^{S-s-1} m\delta^m$.

Figure 1: Cross-Sectional Differences in Investment Time and Wages over the Life Cycle



early in life (left panel), which ranges from virtually no investment for very low-ability individuals all the way to full-time investment for those with highest ability (who enroll in college). This significant cross-sectional heterogeneity is at the heart of the mechanism that generates many of the results in this paper, including the decline in the college premium in the 1970s and the small rise in consumption inequality, among others.

In the right panel, the wage profiles show significant fanning out over the life cycle, which is the only source of wage inequality in this model. The implication that systematic differences in growth rates are the major driving force behind the rise in wage inequality over the life cycle is supported by recent empirical studies that estimate wage and labor earnings processes from micro data sets (Baker (1997), Guvenen (2007, 2009), Huggett, Ventura and Yaron (2006a,b)).¹⁴ For example, the calibrated version of the present model in the next section implies that the cross-sectional wage inequality at age 55 is about 9 times the inequality at age 35. For comparison, Guvenen (2009, table 2) reports that the component of wage inequality that is due to *systematic* differences in growth rates (that is, net of the inequality due to idiosyncratic shocks) is about 10.5 times the inequality at age 35.

¹⁴This approach is different, for example, from the strategy followed by Heckman, Lochner, and Taber (1998) who also allow for ability differences (in the standard Ben-Porath model) but proxy learning ability with the Armed Forces Qualification Test (AFQT) score, which results in very small differences in cross-sectional investment and wage profiles. Compare, for example, figure 1 here to figure 3 in that paper.

3 Quantitative Analysis

In this section, we calibrate the three different versions of the general framework described in the previous section to the U.S. data under the assumption that the U.S. economy was in steady state before SBTC took effect in 1970. The model is solved numerically, and the results below are computed using simulated data. For the stochastic versions (Models 2 and 3 below), we simulate 300 paths for aggregate skill prices and 600 individuals in each cohort. We then compare the evolution of the wage distribution implied by the model from 1970 to 2000 to the data. The U.S. wage data used in this paper are from the annual March Current Population Surveys (CPS) on full-time full-year male workers covering the period 1963 to 2003, and have been provided to us by David Autor; they are the same as the data used in Autor, Katz, and Kearney (2005a,b).¹⁵

Model 1: *Deterministic Baseline (DB)*

The first model is obtained by assuming (i) $\sigma_\varepsilon^2 = 0$, so that there are no shocks to skill prices, and $\eta_0^n \equiv 0$ for all individuals, so that individuals have perfect foresight about the future.

Model 2: *Bayesian Learning with Unbiased Priors (BL⁰)*

The second model is obtained by assuming (i) $\sigma_\varepsilon^2 > 0$, so that there are shocks to skill prices, and individuals learn about the future evolution of skill prices after SBTC in a Bayesian fashion as described above. We consider two versions of this model. We first consider the case when the initial signal is unbiased: $E(\eta_0^n) = 0$ (which is the case described above). We refer to this version with unbiased priors as the BL^0 model.

Model 3: *Bayesian Learning with Pessimistic Priors (BL⁻)*

It is also useful to consider a third case where the initial signal is pessimistic, so that the average forecast of the growth of skill prices is well below the truth. In particular, we assume $E(\eta_0^n) = -\kappa^*$ so that the average initial forecast is $\hat{\kappa}_0 = \kappa^* - \kappa^* = 0$. In this case, individuals on average do not realize the advent of SBTC, and in fact half of them forecast that skill prices will fall. Moreover, in the calibration below, we will choose the parameters such that almost all individuals will initially underestimate the true growth rate of skill prices, and learning will be slow during SBTC. We refer to this pessimistic version as the BL^- model.

3.1 Calibration

Except where noted below, the calibration of parameters is common to all three models. Individuals enter the economy at age 20 and retire at 65 ($S = 45$). The net interest rate, r , is set equal to 0.05, and the subjective time discount rate is set to $\beta = 1/(1+r)$.

¹⁵Our focus on male workers is to abstract from the significant changes during this period in females' labor market participation rates, fertility rates, family composition, and so on, which may potentially have affected the determination of females' wages (more so than males' wages). Having said that, it should be noted that many of the trends documented below have also been observed in females' wages, suggesting that the human capital channels emphasized in this paper could also have played an important role in the evolution of wages for that group as well.

Aggregate Production Function. The growth rate of neutral technology level, Z , is set equal to 1.5 percent per year. As will become clear below, *measured* TFP growth will be different from this number when the amount of investment on the job changes over time. As mentioned above, the curvature of the aggregate production technology, ρ , is set equal to 1.0. Notice that θ_L and θ_H always appear multiplicatively with raw labor and human capital, so the initial values of these parameters serve only as a normalization (given that H and L are also calibrated below). Therefore, we normalize $\theta_{L,t} + \theta_{H,t} = 1$ and set $\theta_{L,t} = \theta_{H,t} = 0.5$ for all $t < 1970$. We calibrate the change in the skill bias of technology after 1970 below.

Human Capital Accumulation. The estimates of α —the curvature of the human capital accumulation function—typically vary between 0.80 and 0.95 (see, for example, Heckman (1976), and the more recent estimates in Heckman et al. (1998) and Kuruscu (2006)). In Guvenen and Kuruscu (2007) we show, theoretically, that if α is higher than a certain threshold the college premium will fall and average wages will stagnate in the short run after SBTC (consistent with the data). Therefore, here we set $\alpha = 0.80$, a value close to the lower end of this empirically plausible range, to show that the plausible quantitative implications found in the next section do not require an extremely high value of α .¹⁶

Accounting for Idiosyncratic Shocks. For a meaningful comparison of the model to the data, it is important to account for the fact that the model abstracts from idiosyncratic shocks, which are present in the data. To this end, we assume that the logarithm of the observed wage in the data can be written as

$$\log \tilde{w}_{s,t}^j = \log w_{s,t}^j + v_{s,t}^j + \xi_{s,t}^j, \quad (14)$$

where $w_{s,t}^j$ denotes the systematic (or life cycle) component of wages and is given by the baseline human capital model in this paper; $v_{s,t}^j$ represents a first-order autoregressive shock process, and $\xi_{s,t}^j$ is a transitory disturbance with variance σ_ξ^2 . Both the innovation to the AR(1) process and $\xi_{s,t}^j$ are i.i.d. conditional on all individual characteristics (including s and A^j). This specification is similar to the econometric processes for wages commonly used in the literature.¹⁷ The key assumption we make is that the variances of these idiosyncratic shocks have been *stationary* during the period under study.¹⁸ Under this assumption, and

¹⁶We have also experimented with values between 0.60 and 0.95, and found that they had a qualitatively small effect on our results. These results are available from the authors upon request.

¹⁷Notice that here we are adding these shocks to the wages obtained from a deterministic human capital model, rather than solving for the human capital choice in the presence of these shocks. In Guvenen, Kuruscu, and Ozkan (2009), we tackled this issue directly and solved for investment and wages in the presence of multiplicative shocks. While that extension came at considerable computational cost, the basic behavior of the model was not substantively different along the dimensions explored in this paper. Therefore, here we opt for the simpler deterministic specification.

¹⁸While several studies have found the variances of idiosyncratic shocks to have increased during this period (Moffitt and Gottschalk (1994), Meghir and Pistaferri (2004), among others), these studies do *not* account for the possibility that the *dispersion* of wage growth rates could have increased during this time,

letting $var(\cdot)$ denote the cross-sectional variance of a variable, we have

$$var(\log \tilde{w}_{s,t}^j) = var(\log w_{s,t}^j) + \sigma_v^2 + \sigma_\xi^2,$$

where σ_v^2 denotes the cross-sectional variance of the AR(1) process across all age and ability groups. Two points are easily noted from this expression. First, the *level* of the variance of wages in the model needs to be adjusted by $(\sigma_v^2 + \sigma_\xi^2)$ before it can be compared to the data. Second, the *change* over time in the variance of observed wages will mirror that in the systematic component ($\Delta var(\tilde{w}_{s,t}^j) = \Delta var(w_{s,t}^j)$), which allows a direct comparison of the *trend* in the model variances to its empirical counterpart.

Similarly, the implications of the specification in (14) for the first moment of wages can also be easily seen: the average of observed log wages equals that of the systematic component, $E(\log \tilde{w}_{s,t}^j | I) = E(\log w_{s,t}^j | I)$, where I denotes a set of individuals—for example, those in the same age or education group. Therefore, both the level of, and the change in, the first moments of log wages in the model can be directly compared to the data.

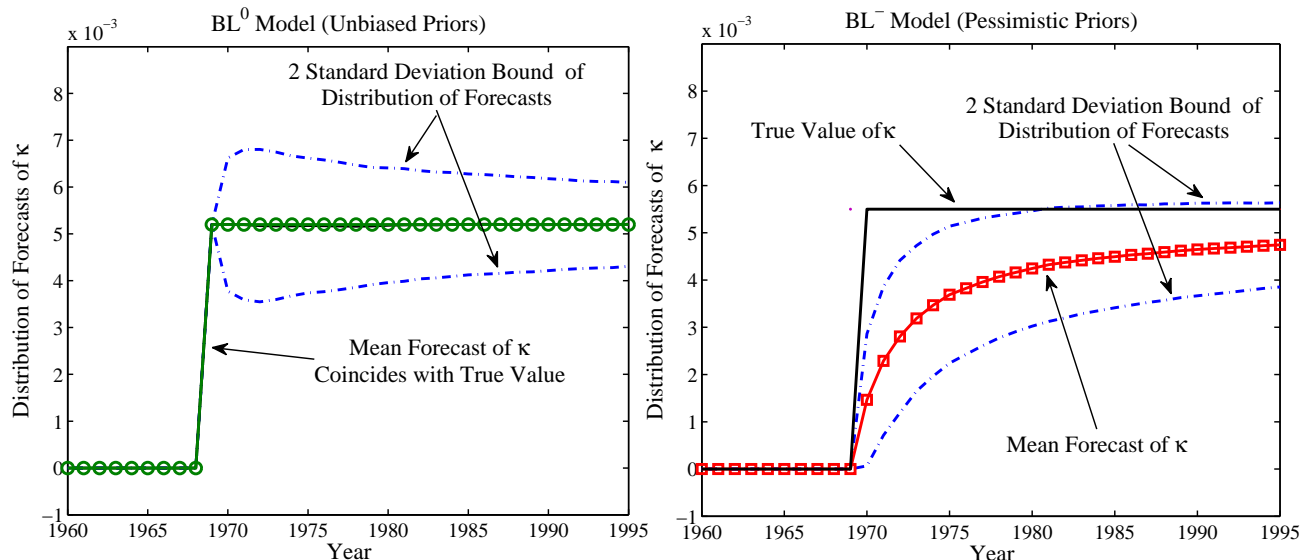
Aggregate Shocks and Prior Beliefs after SBTC. In the stochastic versions of the model (BL^0 and BL^-), the innovation standard deviation of the aggregate shocks to skill bias, σ_ε , is calibrated so that the model is consistent with the variability of the college premium observed in the data. In our data set, the standard deviation of the annual change in the college premium, $\sigma(\Delta\omega_t^*)$, is 1.7 percent per year during the period 1963 to 2003. However, when σ_ε is small, the stochastic versions of the model behave very much like their deterministic counterpart (DB), which is already examined separately. Therefore, we choose a higher target for the volatility of the college premium, $\sigma(\Delta\omega_t^*) = 2.5$ percent, and set $\sigma_\varepsilon = 0.0025$ to match this target. The implied annual volatility of the relative skill prices, $\sigma(\Delta \log(\theta_H/\theta_L))$, is 1.4 percent. Below, we will also examine some implications of the model when skill prices are even more volatile to further illustrate the role of uncertainty and learning on the behavior of the model.

The dispersion of prior beliefs (or initial forecasts of κ^*) is given σ_ε^2/v . A larger value of v reduces the initial heterogeneity in beliefs, as well as slowing down the speed of learning (equation (15) in the Appendix). The goal of the BL versions of the model is to examine how individuals respond to SBTC when they have significant initial uncertainty about SBTC and this uncertainty is not resolved very quickly (otherwise we are back to the DB model). We choose $v = 2.5$, which implies that in the BL^- model, 99.4 percent of individuals initially underestimate the true growth rate of skill prices. This choice also implies a slow rate of learning: 10 years after the start of SBTC, 86.7 percent of individuals still underestimate the true rate of growth of skill prices.

The same information can be seen more clearly in figure 2. The left panel (unbiased priors) plots the true growth rate κ over time as well as the mean forecast (marked with

which is the main thesis of the present paper. Therefore, that evidence is not informative about how the variances should be calibrated over time in our model. Haider (2001, table 4) does allow for dispersion of growth rates to change over time and reports a very small increase in the variance of shocks during this time.

Figure 2: Distribution of Forecasts of Skill Prices ($\widehat{\kappa}_t^n$) Over Time



circles) and the two-standard deviation bounds for the distribution of forecasts (dash-dot lines). The mean forecast overlaps with the true value of κ as expected and the dispersion in beliefs goes down over time, although very slowly. The right panel shows the pessimistic priors case. The mean forecast (marked with squares) starts well below the true value in 1971 and rises gradually; it is still below the truth in 1995, although not substantially. Moreover, the two-standard deviation bounds show that even the most optimistic individuals are underestimating SBTC in this case during the 1970s.

Maximum on-the-job Investment We set χ equal to 0.50, which implies (together with the other parameters below) that in the initial steady state before SBTC, the lowest wage is 51 percent of the average (mean) wage in the economy. For comparison, in 1970 the ratio of the minimum wage to average wage in the US data was 0.49 according to the Organisation for Economic Co-operation and Development (OECD).¹⁹

Distributions of Ability and Raw Labor. Learning ability, A^j , is assumed to be uniformly distributed in the population with the same parameters for every cohort. As for the calibration of individuals' raw labor endowment, note that the present model is interpreted as applying to human capital accumulation after secondary school. But then, the assumption we made in the theoretical model—that individuals start out with the same human capital level—may be too restrictive because it seems likely that different individuals would have accumulated different amounts of human capital by the time they make the college enrollment decision. A simple way to model this heterogeneity is by assuming that the amount of raw labor, l , has a non-degenerate distribution in the population. We also assume l to have a

¹⁹The US data statistic is obtained from the OECD's website (www.stats.oecd.org).

uniform distribution that is the same for all cohorts. Each distribution is fully characterized by two parameters, giving us four parameters to be calibrated.²⁰ The mean value of raw labor, $E[l^j]$, is a scaling parameter and is normalized to one, leaving three parameters: (i) the cross-sectional standard deviation of raw labor, $\sigma(l^j)$, (ii) the mean learning ability, $E[A^j]$, and (iii) the dispersion in the ability to learn, $\sigma(A^j)$. These are chosen to match the following three moments:

1. the average cross-sectional variance of log wages between 1965 and 1969,
2. the average level of the log college premium between 1965 and 1969,
3. the mean log wage growth over the *life cycle*.

As discussed above, we need an estimate of the variances σ_v^2 and σ_ξ^2 to obtain the target value for the cross-sectional wage inequality. Note that, for consistency, these estimates must be obtained from empirical studies that allow for heterogeneity in wage growth rates as implied by the human capital model in this paper.²¹ Guvenen (2009) estimates such a specification and reports σ_ξ^2 to be 0.047. Similarly, σ_v^2 can be calculated to be 0.088 using the estimates in that paper (table 1, row 2). The average cross-sectional variance of log wages in the U.S. data between 1965 and 1969 is 0.239, implying a target value for the first moment in the model ($var(w_{s,t}^j)$) of 0.104. Second, the log college premium in the U.S. data averaged 0.381 between 1965 and 1969 (and does not require any adjustments), which is the second empirical target we choose. Third, and finally, our target for mean log wage growth between ages 20 and 55 is 50 percent for a cohort of individuals who retire before 1970. This number is roughly the middle point of the figures found in studies that estimate life cycle wage and income profiles from panel data sets such as the Panel Study of Income Dynamics (which typically report estimates between 40 and 65 percent; see, for example, Gourinchas and Parker (2002), Davis, Kubler, and Willen (2002), and Guvenen (2007)).

Table 1 displays the implied values for the distributions of A^j and l^j . Notice that the coefficient of variation of ability is more than four times that of raw labor. Overall, heterogeneity in l has a much more modest effect on the quantitative results than does the heterogeneity in ability. Finally, it should be stressed that with this calibration the model also matches the cross-sectional variance of wage growth rates observed in the U.S. data (see the discussion above, in section 2.5).

²⁰Notice that we also need to calibrate the cross-sectional correlation of l and A . Since we interpret the heterogeneity in l as arising from investments made prior to college, and high-ability individuals are likely to have invested more even before college, it seems reasonable to conjecture that A and l will be positively correlated. Indeed, Huggett, Ventura, and Yaron (2006b) estimate the parameters of the standard Ben-Porath model from individual wage data allowing for heterogeneity in A and l , and provide evidence that the two are strongly positively correlated (corr: 0.792). For simplicity, we assume perfect correlation between the two. Furthermore, as will become clear below, the heterogeneity in l implied by our calibration turns out to be quite small, so the choice of perfect correlation is not likely to be critical.

²¹Several well-known papers on income dynamics, such as MaCurdy (1982), Abowd and Card (1989), and Meghir and Pistaferri (2004), restrict wage growth rates to be the same across the population and therefore are not appropriate for calibrating the present model.

Table 1: Baseline Parametrization

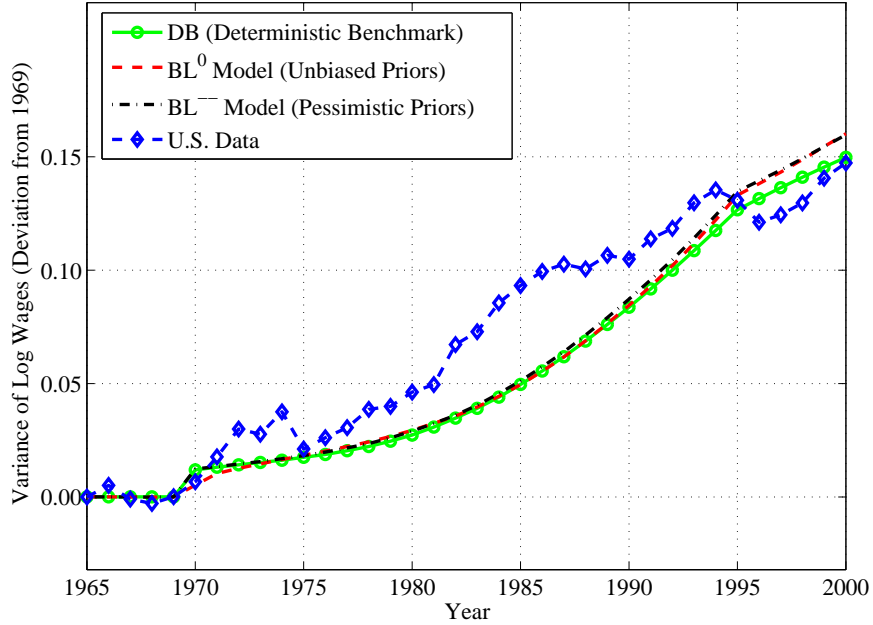
Parameters (common to all versions)		Value	
r	Interest rate	0.05	
β	Time discount rate	$1/(1+r)$	
α	Curvature of human capital function	0.80	
S	Years spent in the labor market	45	
ρ	Curvature of aggregate prod function	1.0	
χ	Maximum investment on the job	0.50	
$\Delta \log Z$	Growth rate of neutral technology	0.015	
T	Duration of SBTC (years)	26	
$E[l^j]$	Average labor endowment (scaling)	1.0	
Parameters: model-specific		DB	BL^0, BL^-
σ_ε	Standard deviation of SBTC shocks	0	.0025
v	Measure of dispersion of initial beliefs	–	2.5
<i>Parameters calibrated to match 1965-69 targets:</i>			
$E[A^j]$	Average ability	.071	0.69
$\sigma(l^j)/E[l^j]$	Coeff. of variation of labor endowment	.0503	.0503
$\sigma[A^j]/E[A^j]$	Coeff. of variation of ability	.245	.242
<i>Parameter calibrated to match 1995 wage inequality:</i>			
$\Delta \log(\theta_H/\theta_L)$	% Annual change in skill-bias (1970-1995)	2.21	2.13, 2.24

Skill-Biased Technical Change. The driving force behind the non-stationary changes in the model is a sustained increase in the relative productivity of human capital relative to raw labor, θ_H/θ_L . Specifically, θ_H grows and θ_L shrinks by κ^* from 1970 until 1995. The ending year is chosen to be consistent with the observation that the rise in wage inequality seems to have slowed down, and productivity growth has started to pick up, by the mid-1990s. This choice is also consistent with empirical evidence indicating a slowdown in the rate of skill-biased technical change in the second half of the 1990s (see, for example, Acemoglu (2002)). However, this choice does not appear to be critical: assuming that SBTC continues until 2010 had very similar implications for the behavior of the model during the 1970s and 1980s.

The main quantitative experiment is the following. After calibrating each version of the model as above, we choose κ^* such that each model matches the overall wage inequality in the U.S. data in 1995 (again, adjusted for the absence of idiosyncratic shocks in the model). The resulting values of κ^* are 0.0054, 0.0052, and 0.0055 in the DB , BL^0 , and BL^- models, respectively. The implied average growth rate of θ_H/θ_L is 2.21, 2.13, and 2.24 percent per year (for a total increase ranging from 70 to 78 percent during the entire period).²² Table 1 summarizes the baseline parameter choices.

²²Notice that, despite the fact that θ_L is falling during SBTC, the *absolute* productivity of raw labor, $Z\theta_L$, continues to grow (by 0.23 percent per year) due to the sustained growth in Z . Therefore, with this calibration, SBTC results only in a *relative* fall in the productivity of raw labor relative to human capital.

Figure 3: The Evolution of Overall Wage Inequality: Model versus U.S. Data, 1965–2000.



4 Model Results

4.1 Evolution of Wage Inequality

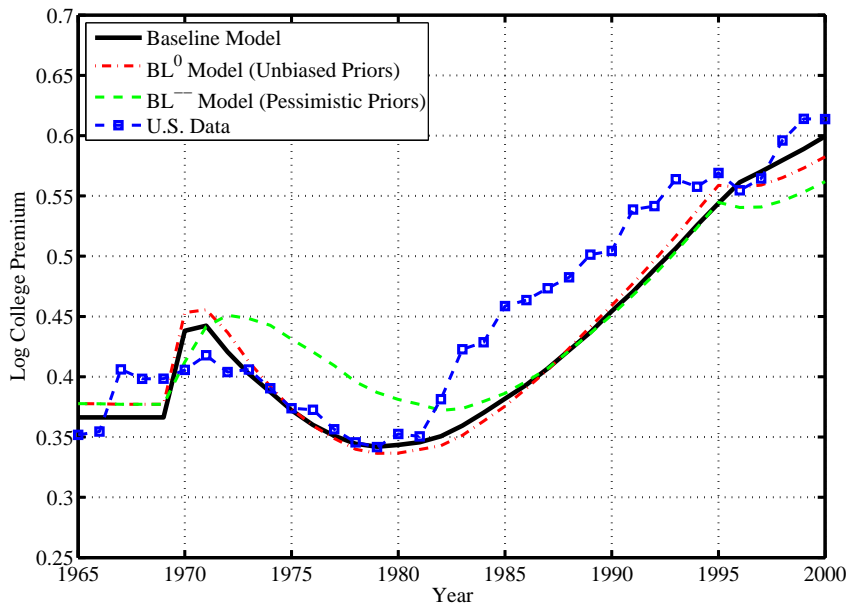
In this section, we examine the implications of the three different versions of our model. As will become clear, the three models have similar implications for several trends and only differ in their implications for certain facts. Therefore, for brevity, in cases where the implications are very similar, we only discuss the intuition for the results in the context of the *DB* model, which is the simplest.

4.1.1 Overall Wage Inequality

We begin with the evolution of overall wage inequality during this period (figure 3). The baseline model is calibrated to match the total change in wage inequality between the first steady state and 1995, and not the evolution between these end points. Yet, the model seems to nicely capture the broad pattern during this period, with a slow increase in the 1970s that accelerates over time.

To understand this convex pattern, two separate effects, which sometimes work in opposite directions, should be noted. First, θ_H/θ_L increases at a roughly linear rate (as can be seen in the left panel of figure 5 below). If there was no change in investment rates in response to SBTC (and therefore, the distribution of human capital remained unchanged over this period), the *price effect* would increase wage inequality at the same constant rate

Figure 4: The Evolution of College Premium: Model versus U.S. Data, 1965–2000.



as the relative price change. However, the investment rate *does* respond to SBTC, which is a key feature of this model. This effect works to offset the price effect early on, because individuals whose investment responds more strongly to SBTC are exactly those with higher ability, and thus who have relatively more human capital already. As a result, the rise in wage inequality is depressed early on. Over time, however, the differential investment response leads to an even larger dispersion in human capital levels, which reinforces the price effect and leads to an accelerating rise in wage inequality. Overall, even though SBTC begins in 1970, most of the rise in overall wage inequality (11.5 out of the 13 log points) happens after 1980, consistent with the U.S. data.

One notable divergence occurs during the 1980s when inequality rises faster in the data compared to the model. Some authors have emphasized the role played by the erosion of the legal minimum wage due to high inflation in the late 1970s, which resulted in the fall of wages in the lower tail of the distribution, thereby increasing inequality (cf., Card and Dinardo (2002)). This factor is not present in the model, which might explain the divergence from the data during the 1980s.

4.1.2 Between-Group Inequality (College Premium)

Figure 4 plots the college premium in the simulated economies along with the empirical counterpart. The DB model is calibrated to match the average level of premium between

1965 and 1969.²³ In the model (thick solid line), the college premium falls throughout the 1970s followed by a robust increase in the next two decades, showing an overall pattern that is both qualitatively and quantitatively consistent with the data.

The Mechanism for the Falling College Premium. Recall that the production function is linear in this model, so the relative supply channel emphasized by Katz and Murphy (1992) is not operational. So how does the model generate the persistent fall in the college premium during the 1970s? To understand this non-monotonic behavior, we decompose the college premium:

$$\omega^* \equiv \frac{\bar{w}_c}{\bar{w}_n} = \frac{(\theta_L L_c^{net} + \theta_H H_c^{net}) / N_c}{(\theta_L L_n^{net} + \theta_H H_n^{net}) / N_n} = \frac{[\theta_L + \theta_H (H_c^{net} / L_c^{net})] (L_c^{net} / N_c)}{[\theta_L + \theta_H (H_n^{net} / L_n^{net})] (L_n^{net} / N_n)}$$

where the subscripts c and n denote college and high school graduates, respectively. N_c and N_n denote the number of college and non-college workers (excluding current students). H_c^{net} and L_c^{net} denote the human capital and raw labor supplied to the market by college graduates. These are calculated as in equation (3) but with the integrals taken over the set of college graduate workers. Other aggregates are defined analogously. So, for example, \bar{w}_c is given by the total wage earnings of all college graduate workers divided by the total number of such workers.

The ratio L_c^{net} / N_c^{net} is approximately equal to the average hours devoted to the labor market (that is, average hours *not* spent on training) by college graduates.²⁴ Divide and multiply the previous equation by θ_L and take logs to get

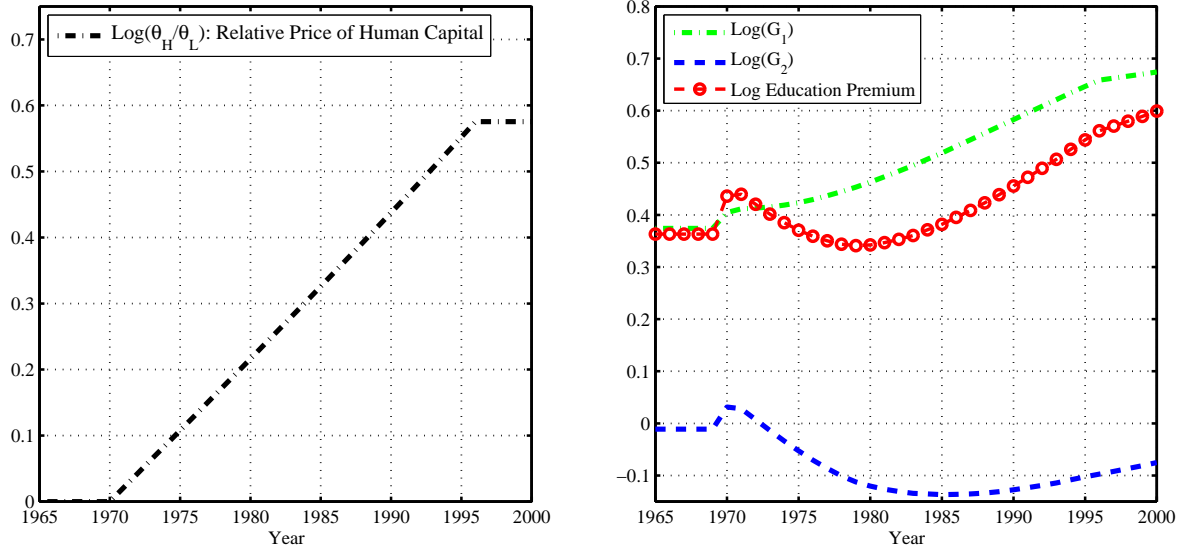
$$\log \omega^* = \log \underbrace{\frac{1 + (\theta_H / \theta_L) (H_c^{net} / L_c^{net}) (L_c^{net} / N_c)}{1 + (\theta_H / \theta_L) (H_n^{net} / L_n^{net}) (L_n^{net} / N_n)}}_{G_1} = \log G_1 + \log G_2.$$

The right panel of figure 5 plots the evolution of the logarithms of G_{1t} and G_{2t} . The term G_{1t} depends on variables that adjust slowly (such as human capital stocks), and it grows monotonically over time. In contrast, there is a steep decline in G_{2t} , especially immediately after SBTC. The reason is that, in response to SBTC, workers with a college degree increase their investment time more than non-college workers (due to the difference in ability between the two groups). Because a rise in investment time reduces L^{net} but has no effect on N (conditional on working), G_{2t} declines significantly in the short run after SBTC. Thus, the log education premium (line with circles in the right panel of figure of 5) initially goes down together with G_{2t} , and over time it bounces back when the decline in G_{2t} tapers off and the

²³A college graduate is defined as an individual who has completed more than two years of full-time investment ($i = 1$). This is analogous to the definition adopted by Autor, Katz, and Kearney (2005a) (i.e., those who enroll in college for more than two years) when constructing the empirical counterpart.

²⁴It can be seen by inspecting equation (3) that if there was no heterogeneity in raw labor, L_c^{net} / N_c^{net} would be exactly equal to the average hours devoted to the labor market. Because the heterogeneity in raw labor is small, this still holds approximately.

Figure 5: Decomposing the College Premium



growth in G_{1t} begins to dominate. Therefore, the differential investment response captured by G_{2t} is crucial for the decline of college premium in the short run.

4.1.3 College Premium with Imperfect Foresight

Despite the fact that any given individual's uncertainty about skill prices does not affect his investment behavior (equation (13)), the fact that each individual holds different mean beliefs does affect average investment, and therefore the behavior of the college premium. To see this, let us compare the investment decisions of two otherwise identical individuals (i.e., same age and ability level) who only differ in their forecasts, given by $\kappa^* + \eta$ and $\kappa^* - \eta$, respectively. From equation (11), optimal investment is a convex function of marginal benefit (as long as $\alpha > 0.5$), which implies that the average of these two individuals' investment will be higher than if they both forecast κ correctly.²⁵ Furthermore, it can also be shown that such a mean-preserving spread will increase investment more for younger individuals (because b_2/b_1 is larger for these individuals, so the same mean-preserving spread in forecasts of κ^* will create a larger dispersion in the marginal benefit of young individuals), and that it will increase investment more among high-ability individuals. Putting these pieces together, it follows that even *without* SBTC, higher belief heterogeneity alone can result in higher investment among high-ability individuals, leading to a fall in the college premium. As individuals learn over time and beliefs converge to each other, this effect will weaken and the college premium will rise again. We now turn to the quantitative results to see whether

²⁵More precisely, let $Q_{s,t}^j(\hat{\kappa})$ denote the investment level when the current forecast is $\hat{\kappa}$. Then it can be easily shown that $\frac{Q_{s,t}^j(\kappa^*+\eta)+Q_{s,t}^j(\kappa^*-\eta)}{2} > Q_{s,t}^j(\kappa^*)$.

this is an important channel for the behavior of (and especially, for the initial decline in) the college premium after SBTC.

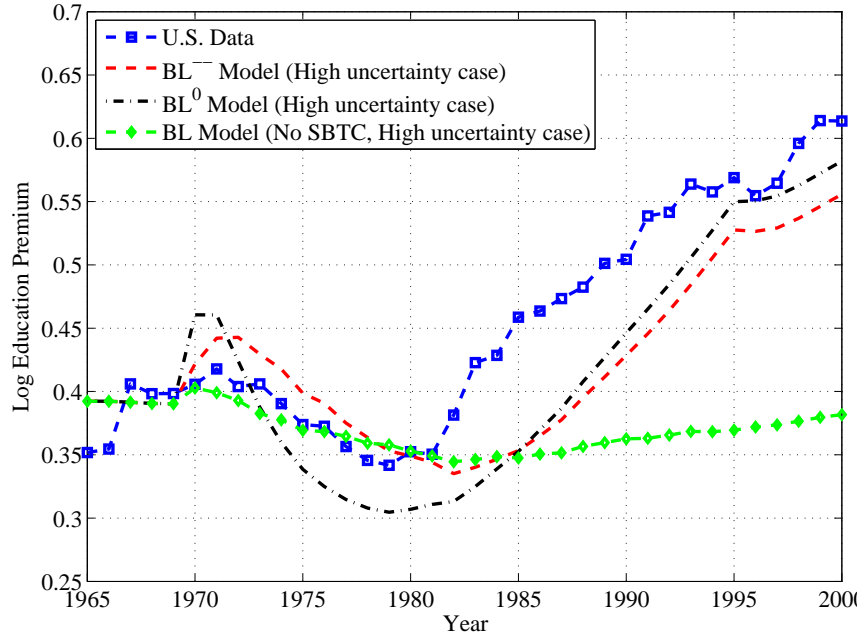
First, in the BL^0 model (dash-dot line in figure 4), the college premium displays a pattern that is qualitatively very similar to the baseline model, but the decline in the premium during the 1970s is somewhat larger (11 log points) compared to the baseline (8 log points). Notice, however, that the baseline model differs from BL^0 in *two* dimensions: in addition to imperfect foresight and Bayesian learning, the latter model also features skill prices that are stochastic, which is not the case in the DB model. In order to disentangle the effects of the two features, we set $v = 10000$, which effectively eliminates all the heterogeneity in beliefs and Bayesian learning from the BL^0 model. We find that the college premium in this case is almost identical to that in the BL^0 model (not shown in the graph), which shows that the heterogeneity in beliefs in the parametrization of the BL^0 model is too small to have a quantitatively significant effect. Thus, the only quantitatively significant difference is that skill prices are stochastic in the BL^0 model, unlike in the deterministic baseline model.

Second, we turn to the BL^- model. Despite the fact that priors are pessimistic and individuals learn slowly, the college premium still falls by 8 log points during the 1970s (dashed line in figure 4). It then rises to reach 54 log points in 1995 (compared to 57 log points in the data). The decline in the college premium even in this case shows that the differential investment channel highlighted in the baseline model is strong enough even when the benefits of investment are underestimated on average. One effect of the pessimistic priors is that the college premium reaches its peak level in 1972 (compared to 1971 in the data) and its bottom in 1982 (instead of 1979 in the data). Given that the choice of 1970 as the beginning year of SBTC is somewhat arbitrary, it is not clear how important these discrepancies are.

It is useful to discuss when, and how, imperfect foresight could affect the college premium in this model. To examine this, we further increase the heterogeneity in initial forecasts (set $\sigma_\varepsilon/\sqrt{v}$ to 50 percent of κ^*), and further slow down the speed of learning ($v = 3.5$). These choices imply $\sigma(\Delta\omega_t^*) = 3.4$ percent per year—double the value in the data. Hence, this case arguably serves as an upper bound on the effect of the lack of perfect foresight. Figure 6 plots the college premium with this new calibration. As can be seen here, the decline in the college premium is more significant now compared to the baseline calibration: it falls by 15.5 log points in the case with unbiased priors (dash-dot line) compared to 11 log points before, and by 10.8 log points in the case with pessimistic priors (dashed line) compared to 8 log points before. The behavior is not affected qualitatively, however.

Finally, to isolate the effect of belief heterogeneity, we set $\kappa^* \equiv 0$ in the last exercise so that there is no SBTC after 1970. Thus, the only thing that happens in 1970 is a pure “belief shock.” Now, the college premium falls by about 5 log points during the 1970s (line with diamonds), after which point it slowly recovers as individuals gradually learn the truth. This result could be anticipated from our previous discussion about the effect of belief heterogeneity on the evolution of the college premium. The conclusion we draw from this exercise is that if the increase in belief heterogeneity is sufficiently large, this alone will result in a decline in the college premium in the short run. However, for parameter values that we

Figure 6: The College Premium with High Uncertainty about Future Skill Prices



consider plausible, this channel does not seem strong enough to be quantitatively important.

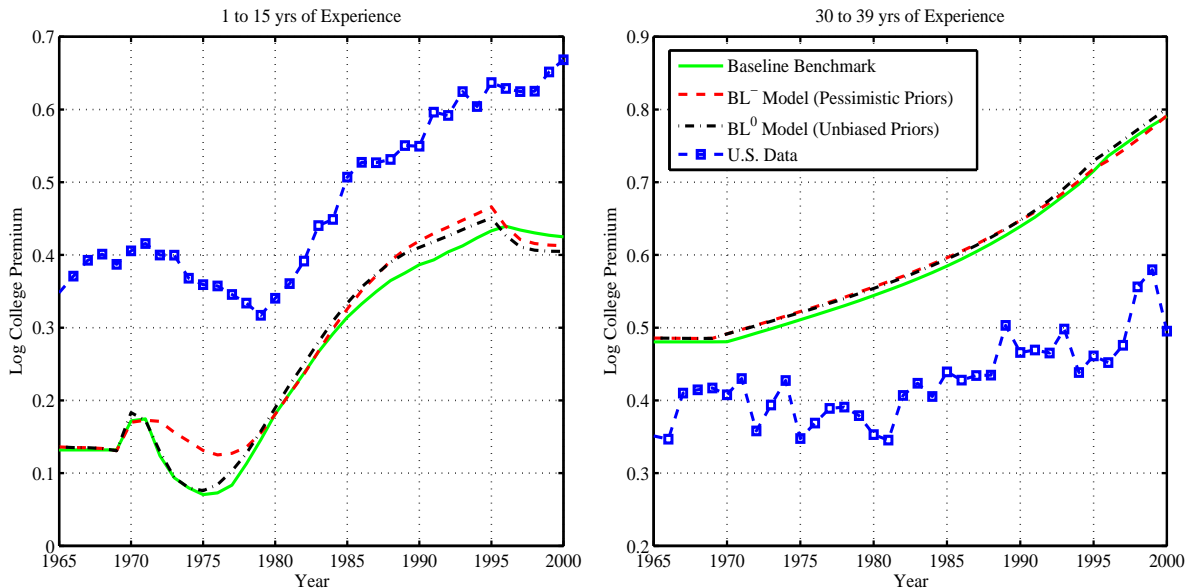
To summarize, these results show that the decline in the college premium in the short run after SBTC does not critically depend on the assumption of perfect foresight. If individuals exhibit moderate heterogeneity in their beliefs immediately after SBTC begins, this has little effect on the college premium. Even when a substantial majority of the population underestimates the true growth rate of the price of human capital for a prolonged period of time, the college premium falls considerably in the short run. Furthermore, if the heterogeneity in beliefs is larger, the decline in the college premium only gets larger.

4.1.4 College Premium within Experience Groups

We now move one step further and examine the college premium *within* different experience groups. Several authors have documented that the fall and rise in the college premium in the U.S. data was largely due to this behavior among young workers, whereas the changes in the college premium among the old was very much muted (Katz and Murphy (1992), Murphy and Welch (1992)). Similarly, Card and Lemieux (2001) have found the same pattern to emerge in British and Canadian data.

In figure 7 we plot the college premium for young workers (1–15 years of experience) and old workers (30–39 years) both in the U.S. data and in the model. For young workers (left panel), the model understates the *level* of college premium before SBTC by about 20 log points. However, the *evolution* of the premium in the model is quite similar to that in the data, with an initial fall of 10 log points in the 1970s, followed by a rise of about 30

Figure 7: The College Premium by Experience Level: Model versus U.S. Data



log points in the subsequent two decades. One difference is that the recovery in the college premium starts in 1977 in the model compared to 1980 in the data. Overall, the evolution of the college premium for young workers in the model is broadly consistent with the data.

Similarly, the right panel plots the college premium among older workers (30-39 years of experience). First, observe that, as noted above, the change in the college premium among these workers is smaller compared to young individuals: the premium does not fall monotonically in the 1970s—the decline averages about 5 log points over the entire decade—and the rise is also smaller—11 log points from 1980 to 1995. In the model (solid line), there is no fall in the college premium in the 1970s; instead it goes up by 5 log points. The premium continues to rise, albeit at a slower pace than among young workers, and increases by 17 log points from 1980 to 1995. Therefore, the model is qualitatively consistent with the fact that the behavior of the college premium during this time is influenced by the large fall and rise in the premium among younger workers and the behavior among the old is less pronounced.

The intuition for these results can be anticipated from the earlier discussions: the large initial fall among young workers is largely due to the fact that these individuals—who face a longer planning horizon, and hence have a larger marginal benefit from investing—respond to SBTC much more strongly than older individuals. In contrast, the slow but monotonic rise in the college premium among older workers is mainly driven by the price effect without a significant investment response. Therefore, the model studied in this paper offers a new, and in our view fairly plausible, explanation for the differences in the behavior of the college premium among different experience groups.

Table 2: Real Wage Changes by Education and Experience Groups, 1971–1987

Group			Change in Log Average Real Wage (multiplied by 100)	
Education	Experience	Sample	1971-79	1979-87
12	Low	Data	0.8	-19.8
12	Low	Model	-1.9	-9.7
12	High	Data	3.2	-2.8
12	High	Model	-1.2	-3.7
16+	Low	Data	-11.3	10.8
16+	Low	Model	-7.8	13.3
16+	High	Data	-4.0	1.8
16+	High	Model	3.8	2.6

Notes: The empirical statistics reported are taken from Katz and Murphy (1992, table 1). The low (high) experience group is defined as workers with 1 to 5 years of experience (26–35 years of experience) in Katz and Murphy (1992) and those with 1 to 15 years of experience (30–45 years of experience) in our model.

4.1.5 Cross-Sectional Wage Profiles by Education and Experience

In documenting the evolution of the relative wages of different groups, labor economists have gone one step further. Since the college premium is essentially the *ratio* of the wages of two groups, further insights can be gained by separately examining the behavior of its numerator and denominator. For example, the aforementioned decline in the college premium among young workers in the 1970s is due to a fall in the numerator (college graduates' wages), whereas the substantial rise in the 1980s is mainly due to a fall in the denominator (high school graduates' wages). Two models that are consistent with the behavior of the ratio can differ greatly in their implications for the evolution of the wages of each education-experience group (i.e., the numerator and denominator), which makes this information useful for distinguishing between alternative explanations.

In table 2, we report the changes over time in the wages of different education-experience groups in the U.S. data (reproduced from Katz and Murphy (1992, table 1)). The most striking fact that emerges from this table happens between 1979 and 1987 (last column). First, among high school graduates, the average wage of workers with few years of experience plummets by 19.8 percent, while older workers only see a small decline of 2.8 percent. As a result, the cross-sectional wage profile of high school graduates significantly *steepens* during this period. Remarkably, the opposite happens among college graduates: young workers see a wage growth of 10.8 percent, whereas older ones only experience a small increase of 1.8 percent. Consequently, the cross-sectional wage profile *flattens* for this group.

We construct the model counterparts of the same statistics with one difference. As we discuss in Section 4.2, the model does not fully capture the magnitude of the slowdown

in average wage growth. Given that our focus here is on the *relative* wage changes across education-experience groups, we normalize the data with the mean wage in a given year before calculating the statistics. This allows us to isolate the relative changes without being distracted by the overstated wage growth for all individuals. The model seems to capture the changes for each education-experience group rather well, not only during the 1980s but also going back to the 1970s. For example, during the 1970s, both in the data and in the model, there is little difference in wage growth by experience levels among high school graduates, whereas for college graduates there is a larger fall for younger individuals than for older ones. More importantly, the model is consistent with the signs and rough magnitudes of wage changes for three of the four education-experience groups from 1979 to 1987 noted above (see the last column of table 2). For the fourth group—young high school graduates—the model implies a significant decline in their relative wages consistent with the data, but it does not capture the full magnitude (9.7 percent in the model versus 19.8 percent in the data). Overall, the average cross-sectional wage profile steepens for individuals with low education and flattens for those with a college degree during this period, consistent with the data.

Three effects drive the wage changes of high school graduates of different ages in the 1980s. *First*, young high school graduates also respond to SBTC (even if it is not to the same extent as college graduates) by increasing their on-the-job investment, which reduces their measured wages. *Second*, there is selection: in response to SBTC, the ability threshold for college enrollment falls, so the average ability pool of high school graduates—those who choose not to enroll in college—also falls, further reducing their wages. Neither one of these channels reduces the wages of *older* high school graduates: since they have a much shorter horizon, they do not increase their on-the-job investment by much, nor do they decide to go back to college to create any compositional change. There is also a *third* effect: young workers have very little human capital, so the main factor they supply is raw labor. Therefore, they suffer from the lower returns to raw labor, but do not benefit from the higher returns to human capital. In contrast, older high school graduates do have some human capital, so they are able to benefit from SBTC, which partly offsets their loss on their raw labor endowment. Put differently, the price effect is negative for the young but close to zero for the old high school graduates. A combination of these three factors, which work in opposite directions for the young and old, explains why the former group experienced a large wage loss while the latter saw no significant change during the 1980s. Notice also that even though SBTC begins in 1970, the three mentioned effects strengthen gradually (as θ_H/θ_L rises) over time, and only begin to make a noticeable impact on the wages of the young much later (1980s).²⁶

The mechanism for the behavior of the wages of college graduates is similar, but the fact that $\chi < 1$ also plays a role. This is because high-ability individuals who want to increase

²⁶One popular explanation offered in the literature for the steepening of the cross-sectional wage profile of high school graduates is that the quality of high school education fell significantly during the 1970s, leading to a fall in the wages of new high school graduates, whereas older workers were already well vested and their wages were protected through union agreements, preventing their wages from falling. The present model generates the same result in a perfectly competitive model, and through a completely different mechanism that is also consistent with other aspects of the evolution of the wage distribution.

their investment significantly in response to SBTC have to stay in college longer due to the upper limit on investment while working. As a result, college students accumulate significant amounts of human capital before entering the labor market. Since SBTC raises the value of human capital, the wages of young college graduates do not fall, unlike those of high school graduates (which can be seen in the right panel of figure 12 below).

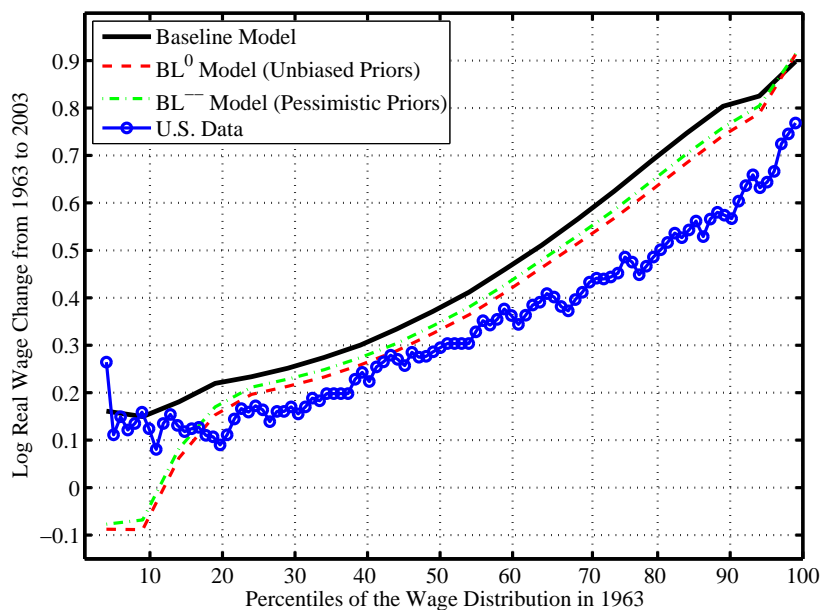
4.1.6 Within-Group Inequality

The analysis so far has focused on the evolution of some key moments of the wage distribution. However, a distribution typically contains much more information than what can be summarized by a few moments, and it is possible for a model to be consistent with some summary statistics, but generate patterns inconsistent with the data at a more disaggregated level. Juhn, Murphy and Pierce (1993) have documented an empirical regularity at a very disaggregated level which presents such a challenge. In figure 8 we report the same finding using our data set which covers a longer time span (solid line). The graph plots how each percentile of the wage distribution in 1963 (horizontal axis) has changed between 1963 and 2003 (vertical axis). The first point to note is that wage growth over this period has been systematically different for *every percentile* of the distribution. This shows that there is more to the rise in overall inequality than can be explained by differences in education alone.²⁷ Second, the relationship between a given percentile in 1963 and wage growth over the subsequent 40 years is almost linear, except at the very low end of the distribution. This implies that wage inequality has increased by a fanning out of the entire distribution, leaving the relative ranking of each percentile largely unchanged over time.

The model counterpart is also plotted in figure 8 (thick solid line). It shows the same general pattern of widening inequality that is spread quite evenly across the wage distribution as observed in the data. Therefore, despite the fact that the model displays significant nonlinearities in the relative wages in the short run, it displays an almost perfect linearity—that is, a stretching out of the entire wage distribution—in the long run. The mechanism behind this result should be clear from earlier discussions. Wage inequality arises entirely from differences in human capital accumulation rates, which in turn arises from differences in ability (for a given age). Because individuals' investment response to SBTC is monotonically increasing in their ability, those with high ability have both higher wages in 1963 and a higher wage growth in the subsequent 40 years; see figure 12 below. The existence of this same pattern in the data suggests that this mechanism appears to be an important channel behind the rise in within-group inequality.

²⁷Juhn, Murphy, and Pierce (1993) also find the same pattern when they examine the wage distribution for each education group and each age group, making this point even stronger. We have generated the corresponding graphs from our model, and they are also qualitatively consistent with the data. For brevity's sake, we do not discuss them here; however, they are available upon request.

Figure 8: Log Real Wage Changes by Percentile: Model versus US Data, 1963–2003



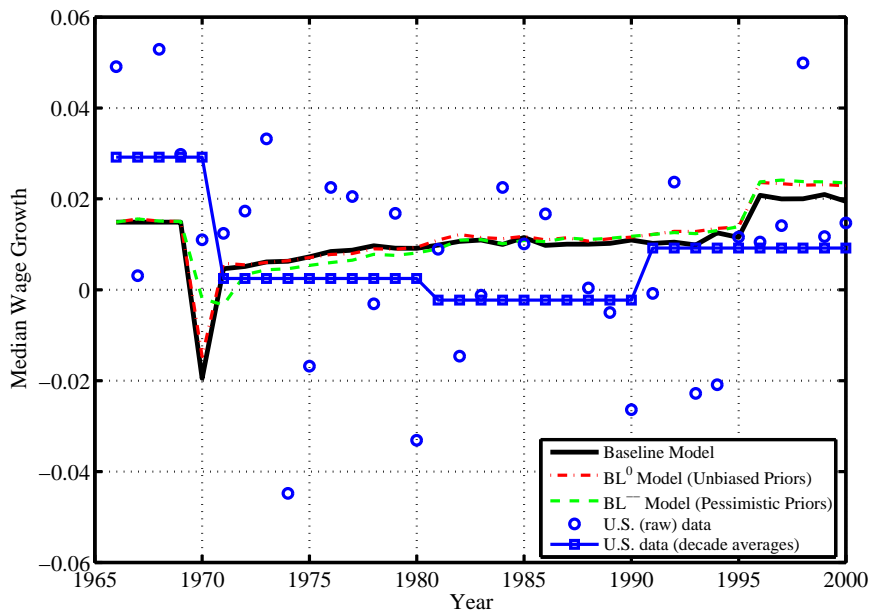
4.2 Evolution of Average Wages

Stagnation of Median Wages and the Productivity Slowdown. We now turn from the second moments of the wage distribution to the first moment, that is, the changes in the average wages over this period. Macroeconomists and labor economists have documented two closely related trends: the slowdown in labor productivity and the stagnation of median wage growth from early 1970s until the mid-1990s.

Figure 9 plots the growth rate of the median wage in the U.S. data and in the model. (The circles correspond to actual U.S. data points, whereas the line marked with squares plots the decade averages to eliminate the short-term fluctuations in the data.) In the US data, median wage growth falls sharply in 1973 and continues to stagnate for several decades: in particular, the median wage grows at 2.6 percent per year in the 1960s, but only by 0.49 percent in the 1970s, and averages close to zero percent until 1995. Median wage growth picks up after 1995 and averages 2.1 percent until the end of the decade. In the model, median wage growth falls sharply in 1970 (immediately after the start of SBTC), and also recovers in a very sluggish manner as in the data: the median wage grows at 0.46 percent per year in the 1970s, and averages 0.81 percent overall until 1995, representing a significant slowdown compared to the 1.5 percent growth during the period before 1970. Thus, while the model does not fully capture the magnitude of the slowdown in median wage growth, it does generate a significant slowdown starting in early the 1970s that persists for several decades. Similarly, figure 10 plots labor productivity growth in the model and in the U.S. data, which reveals a similar picture.²⁸ Finally, as can be seen in both figures, imperfect foresight has a

²⁸More precisely, in the US data, productivity growth averages 2.47 percent from 1951 to 1969, but falls

Figure 9: Growth Rate of Median Wages: Model versus the U.S. Data, 1965-2000

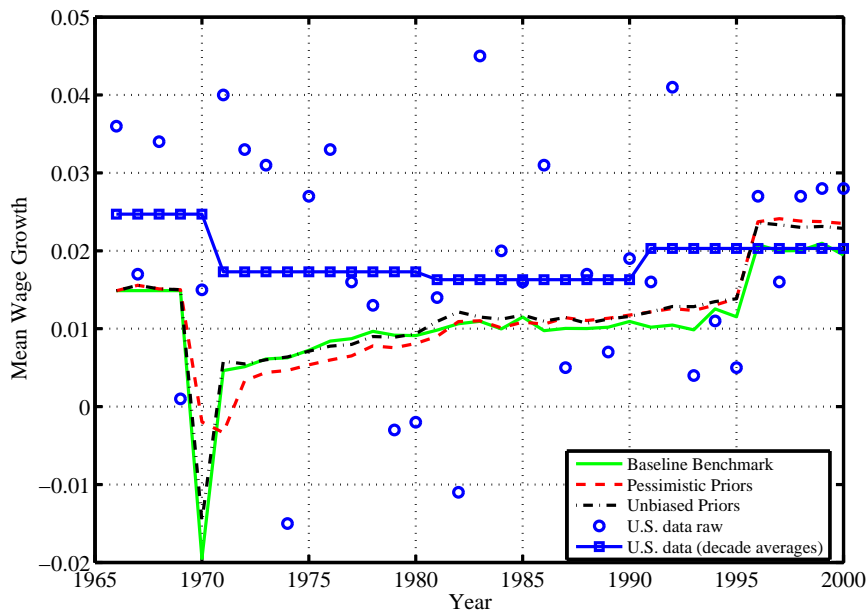


minor effect on the evolution of median wage growth as well as labor productivity.

Three reasons account for the prolonged stagnation in wage growth in the model. *First*, as noted earlier, workers respond to SBTC by increasing their on-the-job investment, which reduces average wages. The fraction of time invested before SBTC is 7.2 percent (or 2.9 hours in a 40-hour workweek) and increases to reach 12.9 percent in 1995 (or 5.1 hours a week). Neither the initial investment level nor the increase during SBTC appears implausibly large, especially considering that what matters for average wages is the change in $(1 - i)$, which goes from 93 percent down to 87 percent over 26 years. One reason for the relatively small change is that, as discussed above, the investment response is concentrated among young individuals, keeping the change in the population average of investment small. Another reason is that on-the-job investment is bounded from above by $\chi = 0.50$, limiting how much it can rise even for young individuals. *Second*, as mentioned earlier, SBTC also lengthens the duration of college education for those already planning to go to college. Consequently, the average ability of individuals who remain in the labor market continually falls during SBTC. Because individuals with lower ability also have low human capital on average, this “selection effect” also depresses average wages and labor productivity after SBTC. *Third*, and finally, there is a pure price effect resulting from SBTC. Essentially, because the price of raw labor is falling as the price of human capital is rising, and because the baseline calibration

to 1.73 percent in the 1970s, and averages 1.65 percent until 1995. Similarly, in the model, productivity growth falls sharply in 1970 and grows by only 0.6 percent per year during the 1970s, but recovers faster and averages 1.24 percent per year until 1995. In the model, labor productivity simply equals the mean wage rate, since there is no capital. The empirical measure of productivity used in this section is calculated (by the authors) from the non-farm output per hour series obtained from the Bureau of Labor Statistic (BLS).

Figure 10: Labor Productivity Growth: Model versus U.S. Data, 1965-2000



implies that the stock of raw labor is larger than the stock of human capital before SBTC, the change in relative prices puts further downward pressure on the average wage.

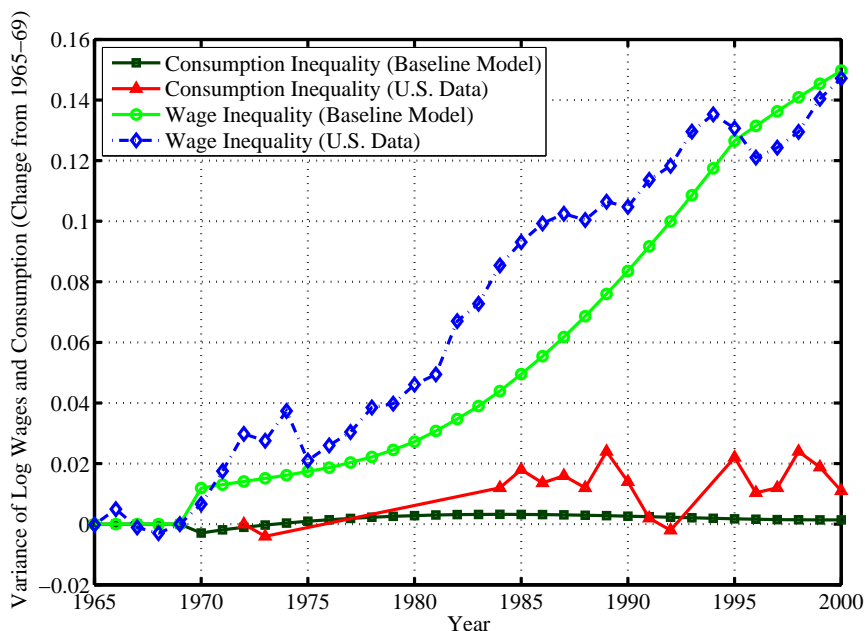
To sum up, during this period the labor market is composed of individuals who invest more on the job, but who also have lower ability than before, resulting in slow wage and productivity growth. Over time, the increase in the total human capital stock due to both types of investment begins to dominate, resulting in a recovery in both the median wage and labor productivity.

4.3 Evolution of Lifetime Wage (Consumption) Inequality

A rather surprising empirical finding from this period is that the rise in consumption inequality has been muted compared to the rise in wage inequality. Figure 11 (line with triangles) plots the variance of log consumption for several years between 1972 and 2000, calculated using the Consumer Expenditure Survey, which shows a very small rise of about 2 log points over this period (data taken directly from Heathcote, Storesletten, and Violante (2005)). Although there remains some disagreement about the exact magnitude of the rise in consumption inequality (mainly due to data problems), other studies have also documented findings broadly supporting this conclusion (cf., Krueger and Perri (2006) and Attanasio, Battistin, and Ichimura (2004)). Moreover, the change between the 90th and 50th percentiles of the consumption distribution has not tracked the large rise in the 90–50 percentile wage inequality. Autor, Katz, and Kearney (2004) document this fact and call it puzzling.

The present model abstracts from many features that would be important for a detailed

Figure 11: The Evolution of Wage and Consumption Inequality: Model versus U.S. Data, 1965–2000.



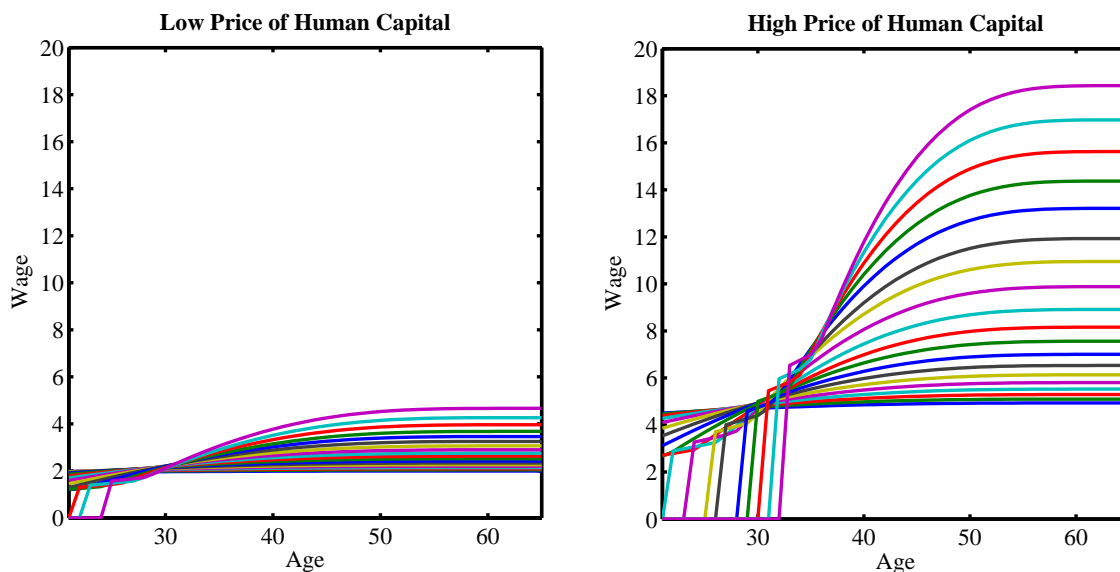
analysis of consumption inequality (such as incomplete markets, retirement savings, demographic changes, etc.). But the model can still address a closely related but key question: Has the substantial rise in cross-sectional wage inequality during this period resulted in a parallel rise in *lifetime* income inequality? The line marked with squares in figure 11 plots the evolution of lifetime income inequality in the model, which shows a very small increase of 0.2 log points during SBTC. Since individuals consume a fixed fraction of their lifetime income in the present model, this is also the rise in consumption inequality.²⁹ (The figure also reproduces the graphs of wage inequality from figure 3 for comparison.)

At first blush, it seems surprising that wage inequality could rise in such a systematic fashion without a significant change in lifetime incomes. But note that, as mentioned earlier, wage inequality rises in this model because of a *fanning out* of wage profiles. This can be clearly seen in figure 12, which plots the life cycle wage profiles in the steady state before SBTC (left panel) and after SBTC (right panel).³⁰ As can be seen in these figures, those

²⁹We sidestep a small complication in this discussion, but we fully account for it in the calculations. In particular, the statement in the text that consumption equals the annuitized value of lifetime income is correct when individuals experience no shocks during their lifetime, since they begin with zero financial wealth. However, for those individuals who are in existence in 1970, consumption after that date is not simply equal to the annuity value of their remaining lifetime income, but also takes into account the fact that they over-saved or under-saved before the shock, given the new path of prices and the implied lifetime wages. Therefore, the inequality in consumption calculated this way goes up more than it would have had we only looked at lifetime income inequality calculated at birth.

³⁰Haider (2001) estimates an econometric process for wages using micro data from 1968 to 1993 allowing for heterogeneity in growth rates. He finds that the rise in wage inequality during this period happened as

Figure 12: Large Rise in *Cross-Sectional* Wage Inequality: Small Rise in *Lifetime* Wage Inequality



individuals who experience a large increase in their wages later in life are exactly those who make larger investments and accept lower wages early on in response to SBTC. Because future gains are discounted compared to the early losses in calculating lifetime income, the rise in lifetime inequality remains small.

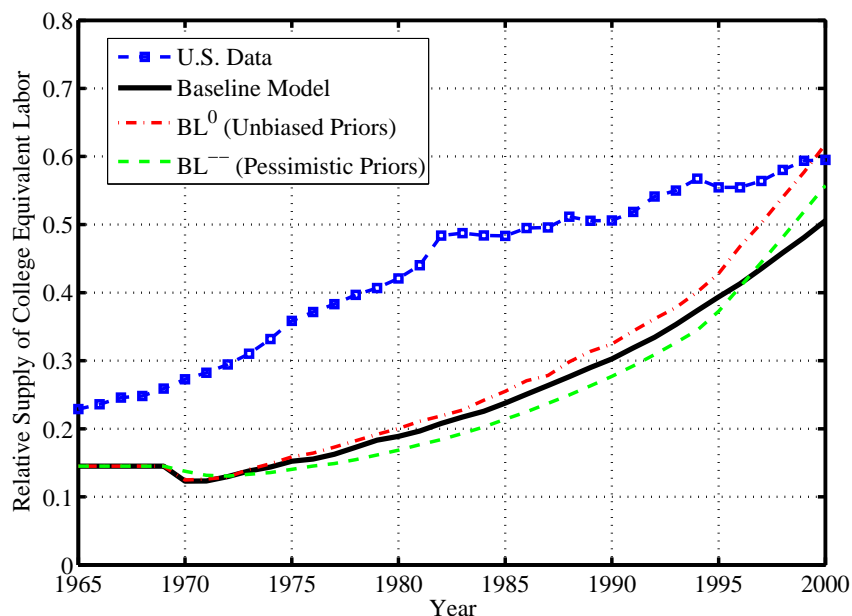
To sum up, the present model offers a new mechanism that is consistent with a large increase in wage inequality but a small change in lifetime inequality. A fuller investigation of this model for consumption facts is left for future work.

4.4 The Rise in the Relative Supply of College Labor

While the main focus of this paper is on the evolution of the wage distribution, the model also makes predictions about the behavior of college enrollment, and consequently, about the change in the relative supply of college-educated labor during this period. Figure 13 plots (line with squares) the total hours worked by individuals with a college-equivalent degree or more relative to those with lower educational levels. This measure more than doubles from 1970 to 2000 in the U.S. data. The model counterpart (thick solid line) understates the level of relative supply before SBTC, which is perhaps not surprising, since no attempt was made to match any aspect of educational attainment in the calibration. However, the relative supply grows significantly in the model, by 0.36 over the entire period, compared to 0.33 in the data. This similarity seems surprising given that in the model college education is modeled merely as a by-product—depending on whether investment exceeds a certain threshold or not—and many potentially important features have been left out, such as tuition costs,

an increase in the *dispersion* of wage growth rates, consistent with this figure.

Figure 13: The Relative Supply of College Equivalent Labor: Model versus U.S. data, 1965–2000



changes in the availability of financial aid for college, changes in the quality of education, and so on. This result suggests that SBTC might have played a more important role than these factors in determining the overall rise in educational attainment during this period.³¹

However, despite these plausible implications for the long-run behavior, the model does not capture the behavior of college enrollment rates in the short run. In particular, college enrollment rates were stagnant in the U.S. data in the 1970s (cf., Card and Lemieux (2001)), whereas the model predicts an immediate rise after the onset of SBTC. As we show in Guvenen and Kuruscu (2007), this counterfactual implication is a direct consequence of the assumption we make in this paper that SBTC happens in a completely *disembodied* fashion: that is, the productivity of all human capital rises at the same rate (given by θ_H) regardless of when it is acquired. Consequently, immediately after the beginning of SBTC, individuals realize that they can gain immensely by investing today and capitalizing on all future improvements in technology, which causes enrollment rates to rise in the short run. In that paper, we also theoretically show that if part of SBTC comes embodied in new vintages of human capital, under certain parameter conditions college enrollment *falls* in the

³¹A related observation during this period has been made by Katz and Murphy (1992, p. 52): “[F]or the 1963-87 period as a whole and most strongly for the 1980s, the groups with the largest increases in relative supplies tended to have the largest increases in relative wages.” This observation is difficult to reconcile with a model where both the demand and supply are driven by exogenous factors. While one explanation for this fact has been provided by Acemoglu’s (1998) model where demand endogenously responds to changes in the supply of different types of labor, the present model is also consistent with this observation, since now both the supply of different types of workers and their wages (due to differential human capital accumulation) respond endogenously to changes in demand.

short run but rises in the long run. This is because, unlike with disembodied SBTC, now individuals do not gain from future improvements in technology by investing in the current vintage of human capital, which makes it optimal to spread investment more evenly over the life cycle. Therefore, investment early in life (i.e., college enrollment) falls, whereas on-the-job investment rises. Because of the latter (and in particular, because on-the-job investment rises in *both* the short run and the long run) the present model’s plausible implications for the short-run behavior of college premium and average wages are preserved. See Guvenen and Kuruscu (2007) for further details and proofs of these results. Although introducing disembodied SBTC would be a very valuable addition for a more detailed study of the trends in educational attainment, such an extension introduces several layers of complexity, which would distract from the main focus of this paper on the evolution of wages. We leave such an extension for future work.³²

5 Conclusion

In this paper we have studied a parsimonious overlapping-generations model of human capital accumulation with some key ingredients: (i) significant heterogeneity in the ability to learn new skills, (ii) raw labor as a factor of production in addition to human capital, (iii) an aggregate production function that takes these two factors—rather than workers with different education levels—as inputs, and (iv) SBTC, modeled as a rise in the relative price of human capital. We have found that the resulting model makes considerable progress toward understanding seemingly disparate trends in the evolution of the wage distribution in a unifying framework.

The model has some other implications for the behavior of wages that for brevity’s sake have not been discussed in the paper. For example, the model is consistent with the flattening of, and the downward shift in, average life cycle wage profiles for subsequent cohorts, documented by Kambourov and Manovskii (2005) from the 1970s to 1990s. Moreover, the model implies that the bulk of the rise in cross-sectional wage inequality happens at the top of the distribution: the rise in the 90–50 percentile wage differential is more than twice the rise in the 50–10 percentile differential, consistent with the finding in Autor, Katz and Kearney (2005b). (Results available upon request). However, the model does not explain why college enrollment was stagnant in the 1970s, which suggests that this model, in its current form, may not be suitable for a detailed study of the educational trends; attempting that as well, however, may be too ambitious a goal for this paper. In Guvenen and Kuruscu

³²There are also other reasons why college enrollment can respond differently to major changes in technology than on-the-job investment. For example, as Becker (1964, p. 51) observes: “Training in new industrial skill is usually first given on the job since firms tend to be the first to be aware of its value, but as demand develops, some of the training shifts to schools.” This is likely to be especially true in the 1970s when many new technologies were developed and used in a decentralized fashion across different firms. It would take time for a systematic body of knowledge to be distilled from the use of these technologies, which can then be taught at formal education institutions. Therefore, one could expect the initial response to SBTC to take place in the form of on-the-job training with a delayed rise in college enrollment.

(2007), we sketch an extension of the present framework where skill-biased technical changes come embodied in new vintages of human capital, which shows promise in that direction.

In contrast to the trend observed in the United States, wage inequality has increased very little in continental European countries (CEU) since the 1970s and, in fact, has fallen in some countries (e.g., France and Spain). At first blush, this observation seems to pose a challenge for the framework developed in this paper, to the extent that one views SBTC as a global phenomenon. In Guvenen, Kuruscu, and Ozkan (2009) we show that the key to understanding these different trends is to take into account the differences in the labor income tax systems between the US and the CEU. In particular, the more progressive tax structures of continental European countries distort the human capital investment decision and mute the response to SBTC, resulting in a small rise in wage inequality over time. We find that a calibrated version of such a model can explain 2/3 of the differences in wage inequality between the US and CEU in 2003, as well as 60% of the differential rise in wage inequality from 1980 to 2005.

Finally, a contribution of this paper that could be of independent interest is the introduction of raw labor, which allows us to think about returns-to-skill in the Ben-Porath framework. The particular specifications for the human capital accumulation function and aggregate production function also make the model quite tractable, and thus suitable for potential extensions.

6 Appendix: Details of Bayesian Learning

6.1 Priors and the Evolution of Beliefs

This section provides the explicit formulas for the Bayesian learning model described in Section 2.3.

Individuals who enter the economy at $t_0 > 0$ observe an initial private signal $\kappa_{t_0}^n = \bar{\kappa}_{t_0} + \eta_{t_0}^n$, where $\bar{\kappa}_{t_0}$ is the average forecast of *existing* individuals in period t_0 , and $\eta_{t_0}^n \sim N\left(0, \sigma_{\kappa, t_0}^2\right)$, where $1/\sigma_{\kappa, t_0}^2$ is the precision of existing individuals at time t_0 . This structure ensures that individuals who are born after the start of SBTC have the same average forecast ($\bar{\kappa}_{t_0}$) and the same precision ($1/\sigma_{\kappa, t_0}^2$) as individuals already in existence in that year.

Every period an individual observes his own wage, which can be written as

$$w_{s,t}^j = Z_t \left[(1 - \theta_{H,t}) l + \theta_{H,t} h_{s,t}^j \right] (1 - i_{s,t}^j),$$

where we used the normalization $\theta_{H,t} + \theta_{L,t} = 1$ made above. Since individuals know the values of all variables except $\theta_{H,t}$, each wage realization reveals the price of human capital (and, consequently, raw labor) in that period. Two consecutive realizations of an individual's wage can then be used to identify $\kappa^* + \varepsilon_{t+1}$ ($= \theta_{H,t+1} - \theta_{H,t}$). Given individual n 's optimal forecast $\hat{\kappa}_{t-1}^n$ at time $t-1$ and his wage realization at time t , his optimal forecast of κ at time t is recursively given by

$$\hat{\kappa}_t^n = \left(\frac{v+t-1}{v+t} \right) \hat{\kappa}_{t-1}^n + \left(\frac{1}{v+t} \right) (\theta_{H,t} - \theta_{H,t-1}), \quad (15)$$

Table 3: Estimating the Katz-Murphy (1992) Regression Using Simulated Data

$\rho = 1$		$\log \omega_t^* = a_0 + a_1 t - (1/\phi) \log(N_{c,t}/N_{n,t}) + error$			
Change from Baseline:		$\hat{\phi}$	R^2	$corr\left(\omega^*, \frac{N_c}{N_n}\right)$	$std(\Delta\omega_t^*)$
(1)	Baseline	2.47	0.88	-0.87	0.025
(2)	$\chi = 0.75$	2.49	0.84	-0.82	0.029
(3)	$N_c/N_n = 0.36$	2.79	0.85	-0.86	0.023

Notes: T=30. The statistics are the medians of 100 simulations.

with precision $(v + t)/\sigma_\varepsilon^2$.

7 Appendix: Estimating the Implied Katz-Murphy Elasticity

There are no existing estimates in the literature of ρ that would guide our calibration. One difficulty with directly estimating ρ from data is that our production function (4) features such inputs as human capital and investment time that are very difficult to measure directly in the data. There is, however, a large literature that has estimated a different elasticity, one that measures the degree of substitutability in a CES production function that takes the labor supplied by college and non-college workers as inputs (cf., Katz and Murphy (1992), Acemoglu (2002)). This elasticity, denoted as ϕ , is obtained by running the following regression:

$$\log \omega_t^* = a_0 + a_1 t - \frac{1}{\phi} \log(N_{c,t}/N_{n,t}) + error, \quad (16)$$

where N_c/N_n is the labor supply of college-educated workers relative to non-college workers, using the notation developed above. Notice that all the variables that appear in this regression can be generated from our model as well. Therefore, an interesting question is the following: what value would we recover for the Katz-Murphy elasticity, $\hat{\phi}$, if we estimate equation (16) using data simulated from our model?

Table 3 reports the results. To obtain these estimates, we solved the stochastic version of the model but abstract from SBTC and, consequently, also from heterogeneity in beliefs (that is, we set $\kappa = 0$ throughout in equation (6), and $\nu \equiv 0$). Then, we simulated data for 30 years and estimated the regression in (16), repeating the exercise 100 times. The reported statistics are the median values from these estimations.

In the first row, the estimated elasticity is 2.47 and the regression has an R^2 of 0.88. The correlation of the college premium and the relative supply is also negative and very large. For comparison, Autor, Katz, and Kearney (2005a) obtain $\hat{\phi} = 2.05$ with an R^2 of 0.94, in the closest specification to ours (see column 3 of table 2 in their paper). Similarly, Hamermesh (1993) surveys the empirical estimates of $\hat{\phi}$ that exist in the literature, which also concentrate around this value.

In the second row, we relax the upper bound on on-the-job investment ($\chi = 0.75$), which may potentially alter the sensitivity of the college premium to the relative supply of college graduates. This change makes little difference to the results. Finally, in the last row, the average relative supply is calibrated to twice its average value in row 1, which increases $\hat{\phi}$ somewhat, to 2.79. Overall, however, these changes have little effect on the estimates of $\hat{\phi}$.

To sum up, the choice of $\rho = 1$, which implies perfect substitution between H and L in our model (in other words, *an infinite* substitution elasticity!), generates not only a finite elasticity between college and non-college workers, but also a relatively small value similar to that observed in the data.

To understand this result, consider the response of this economy to a positive innovation, $\varepsilon > 0$. From equation (6), this shock results in a permanent increase in θ_H/θ_L , and therefore, in a rise in human capital investment. At the extensive margin, this increases the relative supply of college educated workers, starting the year after the shock. At the intensive margin, on-the-job training rises differentially for high- and low-ability individuals, which causes the college premium to fall. Therefore, following a positive ε shock, the supply of college workers rises while, at the same time, the college premium falls. This negative correlation makes it *appear as if* the high supply of college workers reduces the college premium, as would be the case in a CES production function. In other words, the disturbance term in the regression above is positively correlated with the relative supply of college workers, which when ignored biases the estimated coefficient ($1/\phi$) upward (and the estimated elasticity ϕ downward). As a result, our model with $\rho = 1$ generates a finite substitution elasticity in the regression above. Instead, if in the model the supply of college labor were to change for completely exogenous reasons (and the workers to be added were selected randomly from among each group), then the estimate of ($1/\phi$) would be zero, correctly revealing the infinite supply elasticity. We conclude that if skill prices fluctuate in a persistent manner, the estimate of ϕ from the regression in (16) is likely to be downward biased due to an omitted variable bias. The true elasticity may be much higher—in the example presented, it is infinite.

References

- [1] Abowd, John M., and David Card (1989): “On the Covariance Structure of Earnings and Hours Changes,” *Econometrica*, Vol. 57(2), pp. 411-445
- [2] Acemoglu, Daron (1998): “Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality,” *Quarterly Journal of Economics*, Vol. 113, pp. 1055-1089.
- [3] Acemoglu, Daron (2002): “Technical Change, Inequality, and the Labor Market,” *Journal of Economic Literature*, Vol. 40(1), pp. 7-72.
- [4] Attanasio, Orazio, Erich Battistin, and Hidehiko Ichimura (2004): “What Really Happened to Consumption Inequality in the US?” Institute for Fiscal Studies working paper.
- [5] Autor David H., Lawrence F. Katz, and Melissa S. Kearney (2004): “Trends in U.S. Wage Inequality: Re-Assessing the Revisionists,” slides from earlier version of (2005a) paper.

- [6] Autor David H., Lawrence F. Katz, and Melissa S. Kearney (2005a): “Trends in U.S. Wage Inequality: Re-Assessing the Revisionists,” working paper, Massachusetts Institute of Technology.
- [7] Autor David H., Lawrence F. Katz, and Melissa S. Kearney (2005b): “Rising Wage Inequality: The Role of Composition and Prices,” working paper, Massachusetts Institute of Technology.
- [8] Autor, David, Lawrence F. Katz, and Alan B. Krueger (1998): “Computing Inequality: Have Computers Changed the Labor Market?” *Quarterly Journal of Economics*, Vol. (4), 113, pp. 1169-1214.
- [9] Baker, Michael (1997): “Growth-Rate Heterogeneity and the Covariance Structure of Life-Cycle Earnings,” *Journal of Labor Economics*, Vol. 15(2), pp. 338-375.
- [10] Barron, John M., Mark C. Berger, and Dan A. Black (1997): “How Well Do We Measure Training?” *Journal of Labor Economics*, Vol. 15(3), pp. 507-529.
- [11] Becker, Gary S. (1964): “*Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*,” reprinted 1993, University of Chicago Press.
- [12] Ben-Porath, Yoram (1967): “The Production of Human Capital and the Life Cycle of Earnings,” *Journal of Political Economy*, Vol. 75(4), pp. 352-365.
- [13] Berman, Eli, John Bound, and Zvi Griliches (1994): “Changes in the Demand for Skilled Labor within U.S. Manufacturing: Evidence from the Annual Survey of Manufactures.” *Quarterly Journal of Economics*, Vol 109, pp. 367-365.
- [14] Bound, John, and George Johnson (1992): “Changes in the Structure of Wages in the 1980s: An Evaluation of Alternative Explanations,” *American Economic Review*, Vol. 82(3), pp. 371-392.
- [15] Card, David, and John DiNardo (2002): “Skill Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles,” *Journal of Labor Economics*, Vol. 20(4), pp. 733-83.
- [16] Card, David and Thomas Lemieux (2001): “Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis,” *Quarterly Journal of Economics*, Vol. 116(2), pp. 705-46.
- [17] Caselli, Francesco (1999): “Technological Revolutions,” *American Economic Review*, pp. 78–102.
- [18] Davis, Steven J., Felix Kubler, and Paul Willen (2003): “Borrowing Costs and the Demand for Equity Over the Life Cycle,” working Paper, University of Chicago.
- [19] Gourinchas Pierre-Olivier, and Jonathan A. Parker (2002): “Consumption Over the Life Cycle,” *Econometrica*, Vol. 70, pp. 47-89.
- [20] Greenwood, Jeremy (1997): “*The Third Industrial Revolution: Technology, Productivity and Income Inequality*,” AEI Studies on Understanding Economic Inequality, Washington, DC: AEI Press.

- [21] Greenwood, Jeremy, and Mehmet Yorukoglu (1997): "1974," *Carnegie Rochester Conference Series on Public Policy*, Vol. 46(1), pp 49-94.
- [22] Greenwood, Jeremy, and Boyan Jovanovic (1999): "The Information Technology Revolution and the Stock Market," *American Economic Review*, Vol. 89(2), pp 116-121.
- [23] Guvenen, Fatih (2007): "Learning Your Earning: Are Labor Income Shocks Really Very Persistent?" *American Economic Review*, Vol. 97(3), pp. 687-712.
- [24] Guvenen, Fatih (2009): "An Empirical Investigation of Labor Income Processes," *Review of Economic Dynamics*, Vol 12(1), pp 58-79.
- [25] Guvenen, Fatih, and Burhanettin Kuruscu (2007): "Understanding the Evolution of the U.S. Wage Distribution: A Theoretical Analysis," Working Paper No. 13096, National Bureau of Economic Research.
- [26] Guvenen, Fatih, Burhanettin Kuruscu, and Serdar Ozkan (2009): "Taxation of Human Capital and Cross-Country Trends in Wage Inequality," working paper, University of Minnesota.
- [27] Haider Steven J. (2001): "Earnings Instability and Earnings Inequality of Males in the United States: 1967-1991," *Journal of Labor Economics*, Vol. 19(4), pp. 799-836.
- [28] Hamermesh, Daniel S. (1993): "*Labor Demand*," Princeton, NJ: Princeton University Press.
- [29] Heathcote, Jonathan, Kjetil Storesletten, Gianluca Violante (2005): "The Macroeconomic Implications of Rising Wage Inequality in the United States," working Paper, New York University.
- [30] Heckman, James J. (1976): "A Life-Cycle Model of Earnings, Learning, and Consumption," *Journal of Political Economy*, Vol. 84(4), 11-44.
- [31] Heckman, James J., Lance J. Lochner, and Christopher R. Taber, (1998): "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents." *Review of Economic Dynamics*, Vol. 1(1), pp. 1-58.
- [32] Hornstein, Andreas, and Per Krusell (1996): "Can Technology Improvements Cause Productivity Slowdowns?" *NBER Macroeconomics Annual 1996*, Vol. 11, pp. 209-259.
- [33] Huggett, Mark, Gustavo Ventura, and Amir Yaron (2006a): "Human Capital and Earnings Distribution Dynamics," *Journal of Monetary Economics*, Vol. 53(2), pp. 265- 90.
- [34] Huggett, Mark, Gustavo Ventura, and Amir Yaron (2006b): "Sources of Life-Cycle Inequality," working paper, Wharton School, University of Pennsylvania.
- [35] Juhn, Chinhui, Kevin M. Murphy, and Brooks Pierce (1993): "Wage Inequality and the Rise in Returns to Skill," *Journal of Political Economy*, Vol. 101(3) pp. 410-442.
- [36] Kambourov, Gueorgui and Iourii Manovskii (2005): "Accounting for the Changing Life-Cycle Profile of Earnings," working paper, University of Pennsylvania.

- [37] Katz, Lawrence F., and David H. Autor (1999): "Changes in the Wage Structure and Earnings Inequality," in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Vol. 3A, Amsterdam: Elsevier-North Holland, 1463 –1555.
- [38] Katz, Lawrence F., and Kevin M. Murphy (1992): "Changes in Relative Wages, 1963-1987: Supply and Demand Factors," *Quarterly Journal of Economics*, Vol 107(1), pp. 35-78.
- [39] Krueger, Dirk, and Fabrizio Perri (2006): "Does Income Inequality Lead to Consumption Inequality? Evidence and Theory," *Review of Economic Studies*, Vol. 73(1), 163-193.
- [40] Kuruscu, Burhanettin (2006): "Training and Lifetime Income," *American Economic Review*, Vol. 96(3), pp. 832-846.
- [41] MaCurdy, Thomas E. (1982): "The Use of Time Series Processes to Model the Error Structure of Earnings in a Longitudinal Data Analysis," *Journal of Econometrics*, Vol. 18(1), pp 83-114.
- [42] Machin, Stephen, and John Van Reenen (1998): "Technology and Changes in Skill Structure: Evidence from Seven OECD Countries," *Quarterly Journal of Economics*, Vol. 113(4), pp. 1215-44.
- [43] Meghir, Costas, and Luigi Pistaferri (2004): "Income Variance Dynamics and Heterogeneity," *Econometrica*, Vol. 72(1), pp. 1-32.
- [44] Moffitt, Robert A., and Peter Gottschalk (1994): "Trends in the Autocovariance Structure of Earnings in the US, 1969-1987," working paper, Brown University.
- [45] Murphy, Kevin M., and Finis Welch (1992): "The Structure of Wages," *Quarterly Journal of Economics*, Vol. 107(1), pp. 285-326.
- [46] Violante, Giovanni L. (2002): "Technological Acceleration, Skill Transferability, and the Rise in Residual Inequality," *Quarterly Journal of Economics*, Vol. 117(1), pp 297-338.