Ben-Porath Meets Skill-Biased Technical Change: A Quantitative Analysis of Rising Inequality*

Fatih Guvenen† Burhanettin Kuruscu‡

September 19, 2006

Preliminary and Incomplete. Under Heavy Construction.

Abstract

In this paper we quantitatively study the implications of an overlapping-generations model of human capital accumulation for the evolution of the U.S. wage distribution from 1970 to 2000. The 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 early 1970's. Due to the heterogeneity in ability and age, the response of different individuals to SBTC turns out to be systematically different from each other. As a result, the model generates rich behavior in the relative wages of individuals depending on their age and ability. 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. We show that many of our results are not affected by imperfect foresight and indeed some of them become stronger. Overall, the model is quantitatively consistent with several trends including the rise in overall inequality; the fall and rise in the college premium; the rise in within-group inequality; stagnation in median wage growth, and a small rise in consumption inequality despite the large rise in wage inequality.

*For helpful comments and suggestions, we thank Satyajit Chatterjee, Dean Corbae, Tom Sargent, Rob Shimer, Kjetil Storesletten, Aleh Tsyvinski, Gustavo Ventura, Gianluca Violante, Amir Yaron, and especially David Autor and Steve Davis, as well as the seminar participants at the European University Institute, Goethe University Frankfurt, HEC Lausanne, Humboldt University, London School of Economics, Philadelphia FED, Rice University, University of Michigan, University of Rochester, University of Toulouse, UT-Austin, University of Zurich, NBER's Economic Fluctuations and Growth Meeting in Cambridge, Minnesota Macro Workshop, the "Labor Supply and Life-Cycle Productivity" Workshop organized by the Bank of Canada, and the SED conference. We are especially grateful to Daron Acemoglu for comments and suggestions that substantially improved the paper. David Autor kindly provided the data on U.S. male wages used in the paper. Guvenen acknowledges financial support from the National Science Foundation (SES-0351001). The usual disclaimer applies.

†Email: guvenen@eco.utexas.edu; www.eco.utexas.edu/faculty/guvenen

‡Email: kuruscu@eco.utexas.edu; www.eco.utexas.edu/~kuruscu

1 Introduction

[To be Written]

2 A General Framework

We begin by describing a general framework with three key features: (i) human capital accumulation building on Ben-Porath (1967), (ii) a CES aggregate production function that allows imperfect substitutability between the factors of production, and (iii) Skill-biased technical change, and Bayesian learning about the future evolution of skill prices. Then, in the quantitative analysis we consider three benchmarks that are obtained as special cases of this general framework.

2.1 Human Capital Accumulation Decision

The economy consists of overlapping generations of individuals who live for S years. 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 raw labor 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 $h^j_{s,t}$ 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 (potentially stochastic) wage rates, denoted by $P_{L,t}$ and $P_{H,t}$, respectively.

Individuals begin their life with zero human capital, and accumulate human capital according to the following technology:

$$h^j_{s+1,t+1} = (1-\varphi)\, h^j_{s,t} + \underbrace{A^j((\theta_{L,t} l + \theta_{H,t} h^j_{s,t}) i^j_{s,t})^{\alpha}}_{Q^j_{s,t}} \tag{1}$$

where φ is the depreciation rate of existing human capital; $i^j_{s,t}$ is the fraction of time devoted to human capital investment, henceforth referred to as "investment time"; and $Q^j_{s,t}$ 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.[1] 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 out 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 ($i^j_{s,t} = 1$) or none at all in a given period.

The main difference between the Ben-Porath (1967) model and the formulation in (1) is the introduction of raw labor as an additional factor into our model. When $l \equiv 0$ (and $\theta_{H,t}$ is normalized to 1), this model reduces to the standard Ben-Porath model. As will be clear in the analysis below, the reason for our deviation from the standard Ben-Porath model is because it is difficult to sensibly think about SBTC when there is a single skill type.

Investment in human capital takes place on-the-job as long as it does not exceed a fraction χ of an individual's time. If the individual wants to invest more, he enrolls in college and invests 100 percent of his time. Therefore, the choice set for investment time is: $i^j_{s,t} \in [0, \chi] \cup \{1\}$. 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, supplies, etc.), or due to minimum wage laws.[2]

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 will be completely borne by workers, and firms will adjust the hourly wage rate downward by the fraction of time invested on the job (Becker (1965)). Then, the observed wage income of an individual is given by

$$w^j_{s,t} = \underbrace{\left[P_{L,t} l + P_{H,t} h^j_{s,t}\right]}_{x^j_{s,t}} (1 - i^j_{s,t}) = \underbrace{x^j_{s,t}}_{\text{Potential earnings}} - \underbrace{x^j_{s,t}(t)\, i^j_{s,t}}_{\text{Cost of investment}},$$

[1] The dependence of aggregate factor prices and weights in the human capital production function on t is to stress that these could be time-varying.

[2] In 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 $x^j_{s,t}$ 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 foregone earnings while individuals are learning new skills. Since labor supply is inelastic, $w^j_{s,t}$ is also just a scaled version of the individual's observed "wage rate."

2.2 CES Aggregate Production Technology

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

$$\begin{aligned} L_t &= \sum_{s=1}^{S} \mu(s) \int_j l \left(1 - i^j_{s,t}\right) dj, \quad \text{and} \\ H_t &= \sum_{s=1}^{S} \mu(s) \int_j h^j_{s,t} \left(1 - i^j_{s,t}\right) dj, \end{aligned} \tag{2}$$

where $\mu(s)$ is the (discrete) measure of s-year-old individuals, and the integrals are thus taken over the distribution of individuals of all types and ages.[3] Notice that H_t and L_t measure the *actual* amounts of each factor used in production (that is, net of the time allocated to human capital investment) to distinguish them from the "total stocks" of these factors defined later below. The aggregate firm uses these two inputs to produce a single good, denoted by Y, according to the familiar CES production function:

$$Y = Z\left([\theta_L L]^\rho + [\theta_H H]^\rho\right)^{1/\rho}, \tag{3}$$

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 in the same way (compare equations (1) and (3)). In the robustness analysis we relax this assumption, and consider different alternatives specifications for the weights in the human capital function. [To be added]

The firm solves a static problem by hiring factors from households every period to maximize its profit: $Y - P_L L - P_H H$. The factor prices corresponding to human capital and raw labor are:

$$\begin{aligned} P_H &= \frac{\partial Y}{\partial H} = Z\theta_H^\rho \left(\theta_L^\rho [H/L]^{-\rho} + \theta_H^\rho\right)^{\frac{1-\rho}{\rho}}, \quad \text{and} \\ P_L &= \frac{\partial Y}{\partial L} = Z\theta_L^\rho \left(\theta_H^\rho [H/L]^{\rho} + \theta_L^\rho\right)^{\frac{1-\rho}{\rho}}. \end{aligned} \tag{4}$$

[3] For the population structure assumed so far, $\mu(s) = 1/S$.

The price of human capital *relative* to raw labor has a simple expression:

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

While the aggregate production function has the same CES form as commonly used in the literature, its inputs are different than what is typically assumed. In most previous work H and L denote the labor supplied by workers with college and high school education respectively. Therefore, a change in the price of H relative to L 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 1970's resulted from the rapid increase in the supply of college-educated workers (c.f., Katz and Murphy (1992), Juhn, Murphy and Pierce (1993)).

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 relative to L affects all individuals differently depending on their ability level as well as their age, which gives rise to rich dynamics in wage inequality. Moreover, as we show below, the college premium is now very different than 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, which has received a lot of attention in the existing literature. To isolate and highlight the role of the mechanism proposed in this paper for the college premium, we make this assumption in some of our benchmarks below.

2.3 SBTC and Bayesian Learning About Skill Prices

Although it is feasible to introduce uncertainty when $\rho < 1$, this complicates the solution of the model considerably. Studying Bayesian learning about skill prices in this environment adds yet another level of complexity. Furthermore, in the next section we argue that $\rho = 1$ is not an empirically implausible benchmark. Therefore, we will study the case with uncertainty and learning only when $\rho = 1$.

We assume that skill-*neutral* technological progress takes place at a constant rate: $Z_{t+1} =$

$(1+g)\, Z_t$.[4] 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*:

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

where $\varepsilon \sim N\left(0, \sigma_{\varepsilon}^2\right)$. There are several points to mention 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 rate 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 non-zero. Putting the two together shows that κ and ε_{t+1} only affect the relative productivity of factors: $\theta_{H,t+1} + \theta_{L,t+1} = \theta_{H,t} + \theta_{L,t} = 1$, where the last equality is a normalization that we make for convenience.

An 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 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 (and will never happen in our simulations).

We assume that 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 a one-time unanticipated jump in the growth rate of human capital's productivity relative to raw labor, i.e., $\kappa = \kappa^* > 0$, for $t > 0$. Therefore SBTC is a regime change in the growth rate of the productivity of human capital relative to raw labor. As with any transitional experiment, a critical issue here is: what do individuals know about the change in κ, and when do they know about it? We entertain two possibilities. In the first benchmark, we assume that as soon as SBTC begins, individuals have perfect foresight about the future paths of $\theta_{H,t}$ and $\theta_{L,t}$ (and we also assume $\varepsilon_{t+1} \equiv 0$). This assumption, while probably too strong, has the main advantage of providing a clear and simple benchmark.

In contrast, there are many ways to deviate from the perfect foresight benchmark, and it is not immediately clear which one is empirically more plausible. For example, if individuals do not perfectly foresee the future growth rate of skill prices, do they nevertheless agree on their incorrect forecast of the future? Or alternatively, and perhaps more plausibly, does the fact that

[4] Although it might seem natural to allow for *stochastic* growth in Z, such shocks affect both the cost and benefit of human capital investment in the first order condition symmetrically (assuming shocks are permanent). As a result, these shocsk do not generate much interesting action, and for simplicity, we abstract from these shocks in what follows.

each individual is uncertain also means that they each hold different beliefs? (In other words, if individuals were able to pool their information, they would discover that SBTC had started. To the extent that this pooling is not feasible, it seems plausible that individuals would have different forecasts of the future.) If beliefs are indeed heterogenous, is the average forecast consistent with the truth (i.e., unbiased)? Or, does the society systematically under- or over-estimate the growth rate of skills immediately after SBTC happens?

In the second benchmark, we model individuals' beliefs after SBTC in a way that allows for several of these possibilities. 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. At the time of the shock, each individual receives a private signal about the future growth rate. In subsequent periods, each individual observes his wage rate and updates his beliefs about the new value 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).

We now describe the learning process more formally. First note that some cohorts of individuals are already in existence at time 0, whereas others will enter the economy after this date. Let n be an index that uniquely identifies every individual in the history of this economy (which can be obtained by interacting ability type j and the year of birth of each individual, t_0). Each individual n who is alive at time 0 observes an initial private signal about the future growth rate, $\kappa_0^n = \kappa^* + \eta_0^n$, where $\eta_0^n \sim N\left(0, \sigma_\varepsilon^2/v\right)$. Therefore, individual n's initial forecast of future growth rate is κ_0^n with a precision of v/σ_ε^2. Similarly, individuals who enter the economy at $t_0 > 0$ observe an initial private signal $\kappa_{t_0}^n = \overline{\kappa}_{t_0} + \eta_{t_0}^n$, where $\overline{\kappa}_{t_0}$ 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 ($\overline{\kappa}_{t_0}$), and the same precision $\left(1/\sigma_{\kappa,t_0}^2\right)$ as individuals already in existence in that year.

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

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

where we substituted $P_{H,t} = \theta_{H,t}$ and $P_{L,t} = \theta_{L,t}$, since $\rho = 1$; and 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 raw labor) in that period. Two consecutive realizations of an individual's wage can then be used to identify $\kappa^* + \varepsilon_{t+1}$ $(= \theta_{H,t+1} - \theta_{H,t})$. Individual n's optimal forecast of κ^* after observing his wage realizations from time t_0 to $t_0 + T$ is:

$$\widehat{\kappa}_T^n = \left(\frac{v}{v+T}\right) \kappa_{t_0}^n + \left(\frac{T}{v+T}\right) \left(\frac{\theta_{H,t_0+T} - \theta_{H,t_0}}{T}\right), \tag{7}$$

with associated precision $(v+T)/\sigma_\varepsilon^2$. By averaging equation (7) over n it can be shown that the average forecast is unbiased (although it clearly moves around over time).

To summarize, the arrival of SBTC brings with it initial heterogeneity in beliefs about the future growth rate of skill prices. This initial heterogeneity seems plausible given that SBTC represents a structural shift (or a regime change), and it is unlikely that all individuals will initially agree on what it entails. As evidence accumulates over time (in the form of observations on the new path of skill prices), individuals' forecasts converge to each other (and to the truth) while the precision of their forecast increases.

Individuals' Lifetime Income Maximization Problem.—Let $V_s^j\left(h^j;\widehat{\kappa},\theta_L,\theta_H\right)$ denote the lifetime income of individual j who is s-years-old with a human capital stock of h, and a current forecast, $\widehat{\kappa}$. Clearly, the dependence of an individual's lifetime income on $\widehat{\kappa}$ comes from the fact that his investment behavior depends on his perception of the future price sequence, $\{\theta_{L,m},\theta_{H,m}\}_{m=t+1}^{T}$, implied by $\widehat{\kappa}$ and current prices (θ_L,θ_H). The income maximization problem of the agent can be written recursively as:

$$V_s^j\left(h_s^j;\widehat{\kappa},\theta_L,\theta_H\right)=\underset{i_s^j}{\mathrm{Max}}\quad\left[(\theta_L l+\theta_H h_s^j)(1-i_s^j)+\left(\frac{1}{1+r}\right)E\left(V_{s+1}^j\left(h_{s+1}^j;\widehat{\kappa}',\theta_L',\theta_H'\right)|\widehat{\kappa},\theta_L,\theta_H\right)\right]$$

$$s.t.$$

$$h_{s+1}^j=(1-\varphi)\,h_s^j+A^j((\theta_L l+\theta_H h_s^j)i_s^j)^\alpha,\qquad\qquad h_0=0,$$

$$V_{S+1}^j\left(h_{S+1}^j;\widehat{\kappa},\theta_L,\theta_H\right)=0,$$

where primes indicate the value of variables in the next time period, and the evolution of factor prices are given by equation (6). Notice that the individual observes the current period prices before making the current period investment decision.

2.4 Optimal Investment Decision

When $\rho=1$, we can explicitly solve for the investment choice, even in the presence of Bayesian learning. We therefore make this assumption here to illustrate the main trade-offs involved in the investment decision.

We first rewrite the problem to simplify the exposition. Using equation (1) the opportunity cost of investing an amount $Q^j_{s,t}$ can be written as:

$$C^j(Q^j_{s,t}) \equiv (\theta_{L,t}l + \theta_{H,t}h^j_{s,t})i^j_{s,t} = \left(\frac{Q^j_{s,t}}{A^j}\right)^{1/\alpha}.$$

Using this transformation, the problem of an individual born in period t_0 can be written in sequential form as

$$\max_{\left\{Q^j_{s,t_0+s}\right\}_{s=1}^{S}} \quad E_0\left[\sum_{s=1}^{S}\left(\frac{1}{1+r}\right)^{s-1}\left(\theta_{L,t_0+s}l + \theta_{H,t_0+s}h^j_{s,t_0+s} - C^j(Q^j_{s,t_0+s})\right)\right]$$

subject to

$$h^j_{s+1,t_0+s+1} = h^j_{s,t_0+s} + Q^j_{s,t_0+s}, \quad \text{with } h^j_{0,t_0} = 0.$$

We assume no depreciation ($\varphi = 0$) which simplifies the expressions, but is not critical otherwise. The optimality condition which determines the amount of investment at time t is:

$$C^j(Q^j_{s,t})' = E_t\left\{\frac{\theta_{H,t+1}}{1+r} + \frac{\theta_{H,t+2}}{(1+r)^2} + ... + \frac{\theta_{H,t+S-s-1}}{(1+r)^{S-s-1}}\right\}. \tag{8}$$

The left hand side of this equation is the marginal cost, and the right hand side is the (expected) marginal benefit (MB) of increasing an individual's human capital stock. 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 this optimality condition (8) is that an expected increase in the future price of skill will immediately affect current investment decision because of the forward-looking nature of this equation.

Using the functional form for the cost function, the optimal investment choice can be solved for explicitly:

$$Q^j_{s,t} = \left(A^j\right)^{1/(1-\alpha)}\left[\alpha MB\right]^{\alpha/(1-\alpha)}. \tag{9}$$

This expression shows that: (i) individuals with higher learning ability invest more in human capital; and (ii) the response of investment to a change in MB, (either due to an increase in the $\theta_H(t)$ sequence or a fall in interest rates) is increasing in an individual's ability level: $\partial^2 Q^j_{s,t}/\partial(MB)\,\partial A^j > 0$.

Under the perfect foresight assumption, all future skill prices are known, so the expectations operator drops from the right hand side. To see what happens under Bayesian learning after SBTC, substitute the future prices of human capital using equation (6) to get:

$$C^j(Q^j_{s,t})' = E_t\left\{\frac{(\theta_{H,t}+\kappa^*+\varepsilon_{t+1})}{1+r}+\frac{(\theta_{H,t}+2\kappa^*+\varepsilon_{t+1}+\varepsilon_{t+2})}{(1+r)^2}+...+\frac{\theta_{H,t}+\sum_{m=1}^{S-s-1}(\kappa^*+\varepsilon_{t+m})}{(1+r)^{S-s-1}}\right\}$$

$$\Longrightarrow \quad C^j(Q^j_{s,t})' = b_1\theta_{H,t}+b_2E_t\left[\kappa^*\right]$$

where b_1 and b_2 are some positive constants.[5] Notice that the aggregate shocks to the price of human capital play no role in the optimal investment decision. This results from the random walk structure in levels assumed above, and in the quantitative section we show that it provides a fairly good approximation to more general plausible specifications. The only effect of learning is seen in the fact that the forecast of κ^* enters the investment decision rather than its actual value.

To understand the role of heterogenous beliefs on the average investment in the economy, it is instructive to compare the investment decisions of two otherwise identical individuals (same age and ability level) who only differ in their forecast of κ^*. Suppose that at time t, the first individual's forecast is $\kappa^*+\eta$, whereas the second individual's forecast is $\kappa^*-\eta$. From equation (9), it is clear that optimal investment is a convex function of marginal benefit (as long as $\alpha > 0.5$), which implies that the average of the investment of these two individuals will be higher than their investment if they both forecast κ^* correctly.[6] 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 among 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, one could conjecture that even without SBTC, an increase in the heterogeneity in beliefs alone can result in higher investment among high-ability individuals, leading to an initial 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. In the next section, we examine whether this is a quantitatively important channel for the behavior of (and especially, for the initial decline in) the college premium after SBTC.

[5] $b_1 = \frac{\left(1-(1+r)^{-S+s}\right)}{r}$ and $b_2 = \left(\sum_{m=1}^{S-s-1}(1+r)^{-m}\,m\right)$

[6] Let $Q^j_{s,t}(\widehat{\kappa})$ denote the investment level when current forecast is $\widehat{\kappa}$. Then we have:

$$\frac{Q^j_{s,t}(\kappa^*+\eta)+Q^j_{s,t}(\kappa^*-\eta)}{2Q^j_{s,t}(\kappa^*)} = \frac{\left[(b_1\theta_{H,t}+b_2(\kappa^*+\eta))\right]^{\alpha/(1-\alpha)}}{2\left[(b_1\theta_{H,t}+b_2\kappa^*)\right]^{\alpha/(1-\alpha)}}+\frac{\left[(b_1\theta_{H,t}+b_2(\kappa^*-\eta))\right]^{\alpha/(1-\alpha)}}{2\left[(b_1\theta_{H,t}+b_2\kappa^*)\right]^{\alpha/(1-\alpha)}}$$

$$= 0.5\left[1+\frac{b_2\eta}{(b_1\theta_{H,t}+b_2\kappa^*)}\right]^{\alpha/(1-\alpha)}+0.5\left[1-\frac{b_2\eta}{(b_1\theta_{H,t}+b_2\kappa^*)}\right]^{\alpha/(1-\alpha)} > 1$$

3 Quantitative Analysis

In this section we calibrate three 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. We then examine the behavior of several variables from 1970 to 2000. Each version of the model is solved numerically, and the results below are calculated using simulated data.[7] The U.S. data on male wages used to construct the empirical figures and tables in the rest of the paper have been provided to us by David Autor, and are the same as in Autor, Katz and Kearney (2005a,b).

Model 1: *Deterministic Baseline* **(DB)**

The first model is obtained by assuming (i) perfect substitution, $\rho = 1$, and (ii) $\sigma_\varepsilon^2 = 0$, so that there are no shocks to skill prices, and $\eta^n = 0$ for all individuals, so that individuals have perfect foresight about the future.

Model 2: *Stochastic Baseline with Bayesian Learning and Unbiased Priors* **($\mathbf{SBL^0}$)**

The second model is obtained by assuming (i) perfect substitution, $\rho = 1$, and (ii) $\sigma_\varepsilon^2 > 0$, 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: $E(\eta^n) = 0$ (which is the case described above). We refer to this unbiased version of this model as the SBL^0 model.

Model 3: *Stochastic Baseline with Bayesian Learning and Pessimistic Priors* **($\mathbf{SBL^-}$)**

As an extreme case scenario, we also 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 $E(\eta^n) = -\kappa^*$ so that the average initial forecast is $\widehat{\kappa}_0 = \kappa^* - \kappa^* = 0$. In this case, individuals on average do not realize the advent of SBTC, and 50 percent 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 SBL^- model.

Model 4: *Deterministic Baseline with Imperfect Substitution* **(DIS)**

The fourth and last model is obtained by assuming (i) imperfect substitution, $\rho < 1$, and (ii) $\sigma_\varepsilon^2 = 0$, so that there are no shocks to skill prices, and $\eta^n = 0$ for all individuals, so that individuals have perfect foresight about the future. We present the results of this extension in section 5.

[7] When $\rho = 1$ and $\chi = 1$, the model can be solved in closed-form, even in the presence of aggregate shocks and Bayesian learning after SBTC. However, even in that case, the expressions for optimal investment, observed wages, and so on, are extremely complicated.

3.1 Calibration

Except where noted below, the calibration of parameters are common to all three models. Specifically, individuals enter the labor market at age 20 and retire at 65 ($T = 45$). The interest rate is set equal to 0.05, and the subjective time discount rate is set to $\beta = 1/R$, implying that individuals will choose a constant consumption path over their life-cycle (given the absence of uncertainty and borrowing constraints).

Aggregate Production Function.—The growth rate of neutral technology level, Z, is set equal to 1.5 percent per year. As noted before and will become clear below, *measured* TFP growth will be different than this number when the amount of investment on-the-job changes over time.

We take the curvature of the aggregate production technology to be unity (in the first three models that are studied in this section). The fourth model features imperfect substitution and is calibrated and studied in section 5. 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 going to be 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 skill-bias of technology after 1970 below.

Aggregate Shocks and Priors After SBTC.—In the stochastic versions of the model (SBL^0 and SBL^-), 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. As could be expected (and will become clear below), 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 slightly 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 (7)). The goal of the SBL 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 SBL^- 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.

Human Capital Accumulation.—The estimates of α—the curvature of the human capital accumulation function—typically vary between 0.80 to 0.95 (see for example Heckman (1976), and the more recent estimates in Heckman, Lochner and Taber (1998) and Kuruscu (2006)). In Guvenen and Kuruscu (2006) we theoretically show that if α is higher than a certain threshold, the college premium will fall in the short-run after SBTC. Therefore, here we set $\alpha = 0.80$, a value close to the lower end of this empirically plausible range, to show that our quantitative results do not require an extremely high value of α. We have also experimented with values between 0.75 and 0.95, and found that it had a qualitatively small effect on our results.[8] Finally, the estimates of the depreciation rate of human capital are generally small and imprecise. We set it equal to zero following Heckman et al. (1998).

Accounting for Idiosyncratic Shocks.—For a meaningful comparison of the model to the U.S. data, it is important to account for the fact that while the model does allow for aggregate shocks to wages, it abstracts from idiosyncratic shocks, which are clearly present in the data. To this end, we assume that the logarithm of the observed wage in the data of individual j at age s in year t can be written as

$$\log \widetilde{w}^j_{s,t} = \log \widehat{w}^j_{s,t} + \upsilon^j_{s,t} + \xi^j_{s,t}, \tag{10}$$

where $\widehat{w}^j_{s,t}$ denotes the systematic (or deterministic) component of wages, and is given by the models in this paper; $\upsilon^j_{s,t}$ and $\xi^j_{s,t}$ represent an AR(1) and an *i.i.d.* shock process respectively. This specification is similar to the econometric processes for wages commonly used in the literature.[9] The key assumption we make is that the variances of these idiosyncratic shocks—denoted σ^2_υ and σ^2_ξ—have been *stationary* during the period under study.[10] Under this assumption, and letting $var(\cdot)$ denote the *cross-sectional* variance of a variable (taken over individuals of all types and ages at a point in time), we then have: $var\left(\log \widetilde{w}^j_{s,t}\right) = var\left(\log \widehat{w}^j_{s,t}\right) + \sigma^2_\upsilon + \sigma^2_\xi$. Two points are easily noted from this expression. First, the *level* of the variance of wages in the model should be adjusted by $\left(\sigma^2_\upsilon + \sigma^2_\xi\right)$ before it can be compared to the data. Second, the *change* over time in the variance of

[8] Each time the model-specific parameters described below are recalibrated to be consistent with the wage distribution before SBTC. These results are available upon request.

[9] One caveat of this specification should be noted: Because idiosyncratic shocks are additive with the logarithm of the deterministic component of wages, they are in fact multiplicative with the level. Thus, it can be shown that if individuals take the existence of these shocks into account when they make human capital accumulation decisions, this would lead to a different optimal choice than the present one. Although such a modification is possible and the model could be solved numerically, we do not tackle this potential complication here.

[10] Several empirical studies have found the variances of both transitory and persistent shocks to have increased during the period that we study (among others, Moffitt and Gottschalk (1994), Meghir and Pistaferri (2004)). However, an important point to note is that these studies do not account for the possibility that the dispersion of wage growth rates could have increased during this time, which is the main thesis of the present paper. Moreover, given the goal of this paper, it seems natural to abstract from other sources of rise in wage inequality, such as the increasing variances of shocks, to see how much mileage one can get by the mechanism emphasized in this paper alone. This is the approach we pursue in this paper.

observed wages will mirror that in the deterministic component ($\Delta var\left(\log \widetilde{w}_{s,t}^{j}\right) = \Delta var\left(\log \widehat{w}_{s,t}^{j}\right)$) which allows a direct comparison of the *trend* in the model variances to its empirical counterpart.

Similarly, the implications of the specification in (10) for the first moment of wages can also be seen easily. The average of observed log wages will equal that of the systematic component, $E\left(\log \widetilde{w}_{s,t}^{j}|I\right) = E\left(\log \widehat{w}_{s,t}^{j}|I\right)$, where I denotes a set of individuals—for example, those in the same age or education group. Therefore, both the level and the trend in the first moments of log wages in the model can be directly compared to the data.

We are now ready to calibrate the remaining parameters, which are specific to the present model. We calibrate χ to 0.50, which (together with the other parameters below) implies that in the initial steady state before SBTC, the lowest wage is 51 percent of the average (mean) wage in the economy. This choice of χ is consistent with the minimum wage interpretation, considering that the legal minimum wage was about 50 percent of the average wage from 1950 to 1970 in the U.S. data.[11]

Distribution 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. Second, 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.[12] We also assume l to have a uniform distribution that is the same for all cohorts. Each distribution is fully characterized by two parameters, giving us four parameters to be calibrated. Of course, 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, so it seems plausible to have a positive correlation between A and l. Huggett et al (2005) provide evidence that ability and initial human capital are strongly positively correlated (by estimating the parameters of a Ben-Porath model from individual wage data). For simplicity we assume perfect correlation between the two. Furthermore, it will become clear later that the heterogeneity in l does not play a significant role in this model, implying that the choice of perfect correlation is not

[11] See also Hornstein, Krusell and Violante (2006) for the ratio of average wage to the bottom 1 and 10 percent of the wage distribution in the U.S. This choice of χ generates statistics broadly consistent with the figures they report [Report the numbers].

[12] Alternatively, initial heterogeneity could be introduced through differences in h_{0,t_0} which we assumed to equal zero in the baseline model. We have tried this alternative specification in the DB version of the model, and found that it makes very little difference.

likely to be critical.

Since these parameters are model-specific and are not directly observable we choose them so that the model matches some key moments of the data in the first steady state. First, the mean value of raw labor, $E\left[l^j\right]$, is a scaling parameter and is normalized to one. The remaining three parameters: (i) the cross-sectional variance of raw labor, $\sigma\left(l^j\right)$, (ii) the mean learning ability, $E\left[A^j\right]$, and (iii) the dispersion in the ability to learn, $\sigma\left(A^j\right)$, are chosen to match the following moments of the data in the first steady state (the average values between 1965 and 1969):

1. the cross-sectional wage inequality,
2. the level of the log college premium, and
3. the total 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. These estimates can be obtained from empirical studies, but for consistency, they need to be based on an econometric specification that allows for heterogeneity in income growth rates as implied by the human capital model we study. Guvenen (2005a) estimates such a specification and reports σ_ε^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). Finally, the average cross-sectional variance of observed wages between 1965 and 1969 is 0.239, implying a target value for $var\left(\widehat{w}_{s,t}^j\right)$ of 0.104. Second, the average college premium was 0.381 during the same time, which is another empirical target we choose. Finally, we calibrate the total wage growth over the life-cycle to 65 percent. This figure is broadly consistent with studies that estimate life-cycle wage and income profiles from panel data sets such as the PSID, which report estimates that vary between 40 and 70 percent (c.f., Gourinchas and Parker (2002), Davis, Kubler and Willen (2002), Guvenen (2005a)). One caveat of this choice is that these studies rely on data after 1968 (the first year PSID data is available), which coincides with the SBTC period. However, we are not aware of any study that estimates the life-cycle (*not* cross-sectional) wage profiles using data from earlier periods.

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. In fact, heterogeneity in l has a much more modest effect on the quantitative results than does the heterogeneity in ability.

Skill-Biased Technical Change.—The driving force behind the non-stationary changes in the model is a sustained increase in 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 latter year is chosen to be roughly consistent with the observation that the rise in wage inequality seems to have slowed down

Table 1: Baseline Parameterization

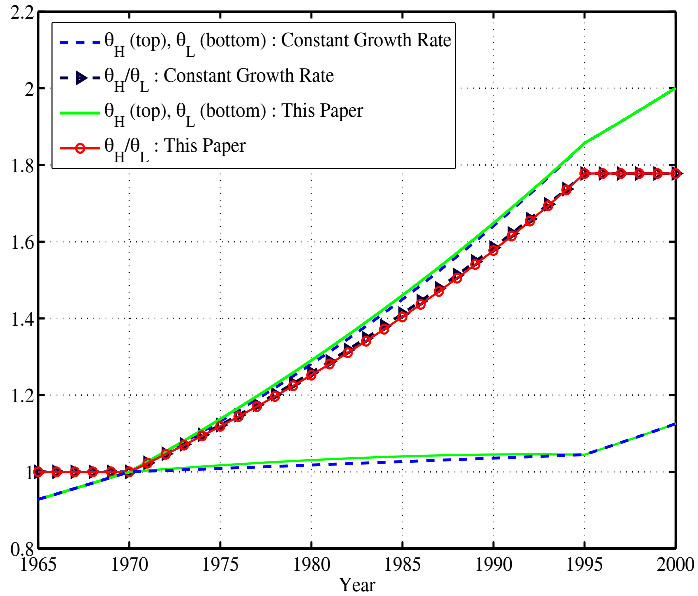
Parameter				Value
R	Interest rate			0.05
β	Time discount rate			$1/R$
α	Curvature of human capital function			0.80
φ	Depreciation rate of human capital			0.0
T	Years spent in the labor market			45
ρ	Curvature of aggregate prod function			1.0
χ	Maximum investment on the job			0.50
$\Delta \log Z$	Growth rate of neutral technology			0.015
σ_ε	Standard deviation of SBTC shocks			.0025
υ	Measure of dispersion of initial beliefs			2.5
$E\left[l^j\right]$	Average labor endowment (scaling)			1.0
Model:		DB	SBL^0	SBL^-
Parameters calibrated to match 1965-69 targets:				
$E\left[A^j\right]$	Average Ability	.071	.069	0.69
$\sigma\left(l^j\right)/E\left[l^j\right]$	Coeff. of variation of Labor endowment	.0503	.0503	.0503
$\sigma\left[A^j\right]/E\left[A^j\right]$	Coeff. of variation of Ability	.0245	.242	.242
Parameter calibrated to match 1995 wage inequality:				
$\Delta \log\left(\theta_H/\theta_L\right)$	Annual change in skill-bias (1970-1995)	2.21%	2.13%	2.24%

by the mid-90's. 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 70's and 80's.

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, SBL^0 and SBL^- 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).[13] As mentioned earlier, we specified the change in the productivities of each factor in their levels (equation (6)), rather than the more common specification in growth rates. For our parameter choices this makes little difference as can be seen in figure 1. The dashed lines plot the paths of $Z\theta_H$ and $Z\theta_L$ assuming a constant growth rate, whereas the solid lines plot the same variables' evolution in our deterministic specification (DB model). Similarly, the line marked with triangles (circles) plots θ_H/θ_L with a constant growth rate (in our baseline). The difference between the two paths is almost negligible, but the level specification simplifies computation significantly (especially in cases with Bayesian learning).

[13] Although it is possible to choose the entire path of θ_H/θ_L during this period to match the path of wage inequality, we do not pursue this approach. Instead, we choose the simplest path—a constant increase in skill bias per year during the transition phase.

Figure 1: Comparing the Paths of θ_H and θ_L during SBTC under Alternative Specifications



Finally, it can also be seen from figure 1 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) during this period 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. Table 1 summarizes the baseline parameter choices.

3.2 Evolution of Wage Inequality

3.2.1 Overall inequality

In the following analysis we present the results from the deterministic benchmark (henceforth called the DB model), the stochastic benchmark with Bayesian learning and unbiased priors (henceforth, the SBL^0 model) and with pessimistic priors (henceforth, the SBL^- model). As will become clear below, 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 will only discuss the intuition for the results in the context of the DB model, which is the simplest.

We begin by analyzing the implications of these models for the evolution of total wage inequality during this period. Figure 2 plots the variance of log wages generated by the DB, SBL^0 and SBL^- models, together with the empirical counterpart. Remember that the models were calibrated to match the levels of wage inequality in 1965-69 and 1995, and not the evolution between these end points. Yet, all three versions of the model seem to nicely capture the broad pattern during this period, with a slow increase in the 1970's that accelerates over time.

To understand the evolution of overall inequality two separate effects, which sometimes work in opposite directions, should be noted. First, the price of human capital relative to raw labor (θ_H/θ_L) increases at a roughly linear rate as can be seen in the left panel of figure 5. If there was no change in investment rates in response to SBTC (and therefore, the distribution of human capital remained unchanged over this period) this price effect would increase wage inequality at the same constant rate 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, most of the rise in overall wage inequality (11.5 out of the 13 log points) happens starting in the 1980's, consistent with the U.S. data.

One notable divergence occurs during the 1980's 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 70's, which resulted in the fall of wages in the lower tail of the distribution, thereby increasing inequality (c.f., Card and Dinardo (2002)). This factor is not present in the model which might explain the divergence from the data during the 80's.

Another point to observe is that inequality continues to increase after 1995 when the price of human capital stops increasing. This is due to the fact that older cohorts with lower dispersion in human capital levels retire from the economy, and are replaced by younger cohorts with higher dispersion. As a result, wage inequality in this model will continue to rise until year 2040, when the population is composed only of individuals born after 1995.

3.2.2 The College Premium

Figure 3 illustrates the behavior of the college premium implied by different versions of the model along with the empirical counterpart. Recall that the only data point targeted in the calibration was the average level of premium between 1965 and 1969. Here, the implications of the three models

Figure 2: The Evolution of Overall Wage Inequality: Model versus U.S. Data, 1965—2000.

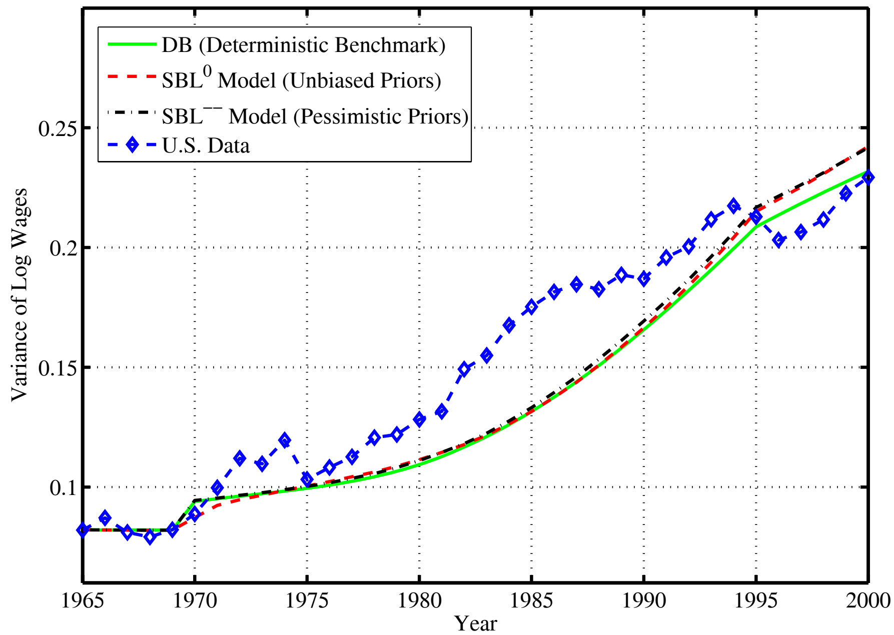


Figure 3: The College Premium: Model versus U.S. Data, 1965-2000.

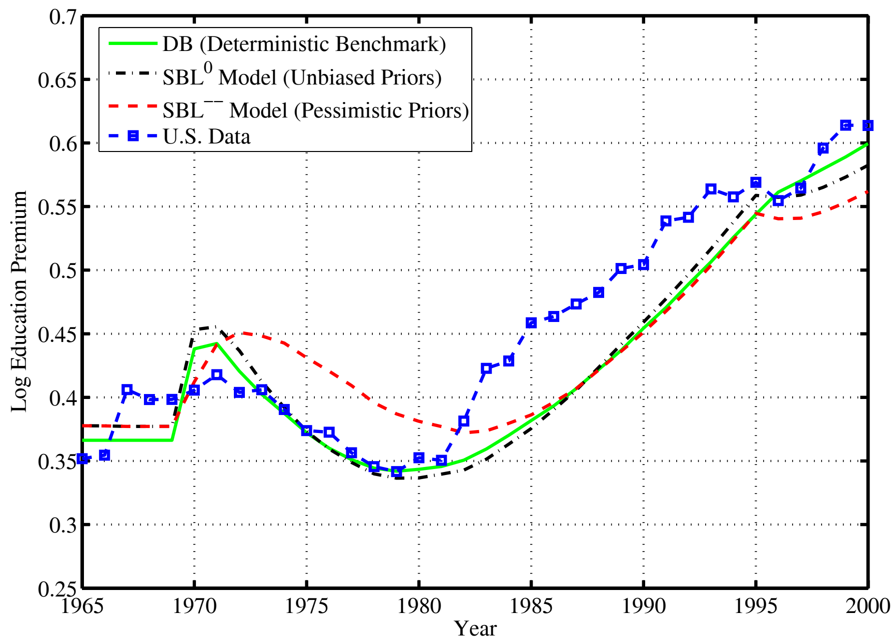
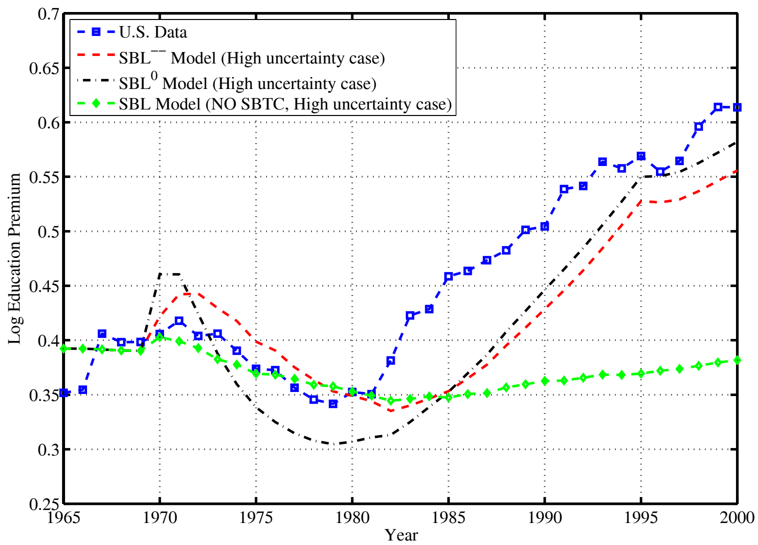


Figure 4: The College Premium with High Uncertainty about Future Skill Prices



are somewhat different from each other, so we discuss them separately.[14]

In the DB model (solid line), the college premium falls throughout the 1970's followed by a robust increase in the next two decades, showing an overall pattern that is both qualitatively and quantitatively consistent with the data.[15] To gain a better understanding of the behavior of the college premium, we use the following decomposition:

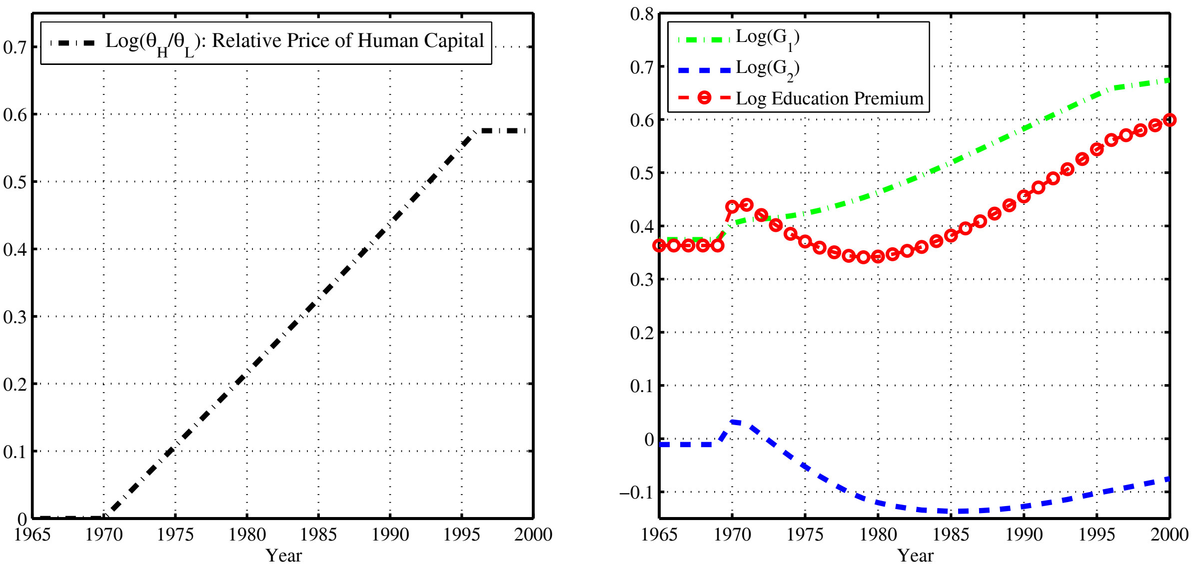
$$\omega^* = \frac{(\theta_H H_c + \theta_L L_c)/N_c}{(\theta_H H_{nc} + \theta_L L_{nc})/N_{nc}} = \frac{[\theta_H (H_c/L_c) + \theta_L]}{[\theta_H (H_{nc}/L_{nc}) + \theta_L]} \frac{(L_c/N_c)}{(L_{nc}/N_{nc})}$$

where $H_c = \sum_{s=1}^{S} \mu(s) \int_{j \in C} h_s^j \left(1 - i_s^j\right) dj$, that is, the human capital supplied to the market by college graduates, and $L_c = \sum_{s=1}^{S} \mu(s) \int_{j \in C} l^j \left(1 - i_s^j\right) dj$. Other aggregates are defined analogously, and the subscript "nc" denotes high-school graduates. Note that we divide both the numerator and denominator by the number of people in that group who are currently active in the labor market: $N_c = \sum_{s=1}^{S} \mu(s) \int_{j \in C} 1\left\{i_s^j < 1\right\} dj$ to get average wages for each type. Note that (L_c/N_c) is equal

[14] **An important point to note is that currently we are not adjusting for compositional changes when constructing the college premium in this section.** This will be done and figures will be updated accordingly.

[15] A natural concern could be whether the falling premium during the 70's is driven by a small number of college graduates who are investing close to 100 percent and therefore receiving wages close to zero. This is not the case, since the fraction of time devoted to investment on the job is bounded from above by $\chi = 0.50$.

Figure 5: Decomposing the College Premium



to the average hours devoted to the labor market (that is, average hours *not* spent on training) by college graduates. Let $H_c/L_c = k_c$, and $H_{nc}/L_{nc} = k_{nc}$. Divide and multiply the previous equation by P_L to get:

$$\log \omega^* = \log \underbrace{\frac{(\theta_H/\theta_L)k_c + 1}{(\theta_H/\theta_L)k_{nc} + 1}}_{G_1} \underbrace{\frac{(L_c/N_c)}{(L_{nc}/N_{nc})}}_{G_2} = \log G_1 + \log G_2.$$

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, 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, i affects L_c negatively, but has no effect on N_c (conditional on working), the rise in investment time causes a decline in G_{2t}, especially early on. Thus, the log education premium (line with circles) initially goes down together with G_{2t}, and over time it bounces back when the growth in G_{1t} begins to dominate.

Turning to the SBL^0 model in figure 3, the college premium shows a qualitatively similar pattern, but the fall in the premium during the 70's is somewhat larger (11 log points) compared to the DB model (8 log points). But overall, the difference between the college premium implied by the DB and SBL^0 models is small. Notice that SBL^0 model does not only differ from the DB model in the imperfect foresight (and heterogeneity in initial forecasts) but also in the fact that there are shocks to skill prices. With this baseline parameter choices it turns out that the heterogeneity in beliefs play a minor role. For example, when $v = 10000$ (which effectively eliminates heterogeneity in priors), the shape of the college premium remains almost identical to that plotted for the SBL^0 model in the figure. This is because the baseline parameterization implies only small differences in initial forecasts after SBTC. Therefore, the (small) difference between the DB and SBL^0 models with the baseline calibration is mainly due to the existence of shocks in the latter model. We explore the role of larger heterogeneity below.

We next turn to the SBL^- model where 99.4 percent of individuals initially underestimate the true growth rate of skill prices. Despite this, the college premium still falls by 8 log points during the 70's, less than in the SBL^0 model but similar to that in the data. It then rises to reach 54 log points in 1995 (compared to 57 log points in the data).[16] Note that the effect of underestimation is very different than an economy where skill prices actually grow slowly: here because individuals underestimate the true growth rate they invest at a lower rate than in the previous models, but skill prices do actually grow fast pushing the college premium upward. Therefore, underestimation of skill prices on its own tends to generate a rise in the college premium. The decline in the college premium even in this case shows that the differential investment channel is quite strong even when the benefits of investment are underestimated. 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). Of course, given that the choice of 1970 as the beginning year of SBTC is somewhat arbitrary it is not clear how important these discrepancies are.

Before closing this section, it is useful to discuss when, and how, imperfect foresight could affect the conclusions we reached about the behavior of the college premium implied by this model. To this end, we increase the initial heterogeneity in forecasts, and further slow down the speed of learning. Specifically, we choose the parameter values such that the standard deviation of the initial forecast, $\sigma_\varepsilon/\sqrt{v}$, is 50 percent of κ^*, and set $v = 3.5$. (These choices also imply that the standard deviation of the change in the college premium is 3.4 percent per year—double that in the data. So in a sense this case arguably serves as an upper bound).

Figure 4 plots the behavior of the college premium (with this new calibration) when the prior beliefs are unbiased (dash-dot line) and when they are pessimistic (dashed line). As can be seen here, the decline in the college premium is more significant now compared to the baseline calibration:

[16] Note that if the news about SBTC arrived in the 1970's, but the actual increase in the price of human capital started later this would make the fall in the college premium more prominent in the 1970's (because investment and hence foregone earnings would rise in the 1970's for college educated workers, but the benefits of a high skill price would still not be realized). Therefore, to the extent that the rise in skill prices was not linear as we assumed so far, but happened in a more convex fashion over time—a slow start that accelerated over time—it would only reinforce the mechanism described here.

it falls by 15.5 log points in the former case (compared to 11 log points before), and by 10.8 log points in the latter (compared to 8 log points before). The behavior is not affected qualitatively, however.

To isolate the effect of heterogeneity in beliefs, we also consider the effect of a pure "belief shock." Specifically, suppose that $\kappa^* = 0$, so that there is no SBTC after 1970. The only change that happens in 1970 is that each individual observes a private signal (as before) with zero mean about the future growth of skill prices. Over time then beliefs converge to the true value of zero. The line marked with diamonds shows the behavior of the college premium in response to this pure belief shock: the college premium falls by about 5 log points during the 70's, after which point it slowly recovers as individuals gradually learn the truth. This result could be anticipated from our previous discussion in section 2.4 about the effect of belief heterogeneity on the evolution of the college premium. This last exercise show that if the rise in belief heterogeneity is sufficiently large, this in fact strengthens the mechanism that causes a decline in the college premium in the short-run. However, for parameter values that we consider plausible, this channel does not seem to be large enough to be quantitatively important.

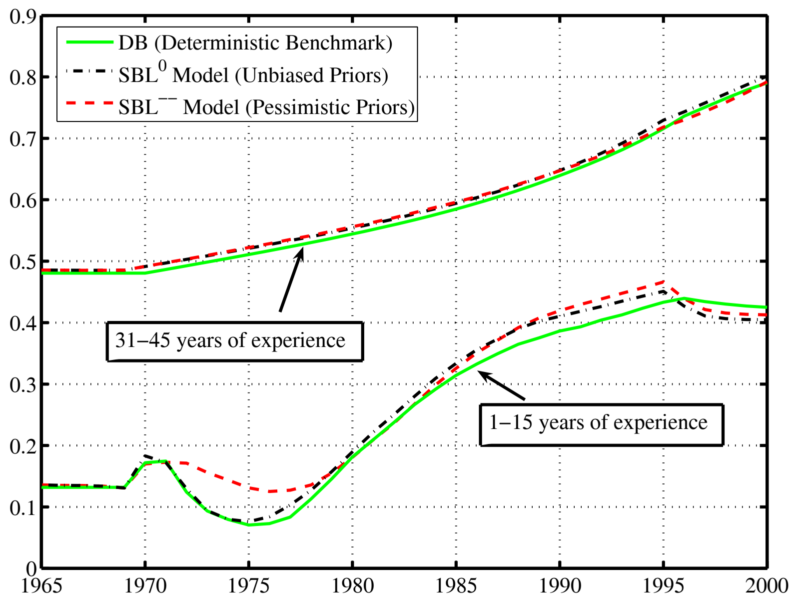
To summarize, these results show that the decline in the college premium in the short-run after SBTC does not critically depend on the assumption of perfect foresight. If individuals exhibit moderate heterogeneity in their beliefs immediately after SBTC begins, this has little effect on the behavior of 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.

3.2.3 The College Premium within Experience Groups

A well-documented fact is that the behavior of the college premium in the U.S. during this period has been different for different experience groups (Katz and Murphy (1992), Murphy and Welch (1992), Autor, Katz and Kearney (2005)). In particular, these authors show that the fall and rise in the overall college premium discussed in the previous section was largely attributable to this behavior among individuals with less experience. In contrast, the fall and rise in the premium among more experienced individuals has been very much muted. Similarly, Card and Lemieux (2001) focus on age-groups (rather than experience), and examine data from U.K. and Canada in addition to the U.S. They find the same pattern to emerge in these countries as well.

In figure 6 we plot the college premium for two different experience groups (1-15 and 30-45) implied by the DB model. The college premium is higher among more experienced individuals

Figure 6: The College Premium By Experience Level in the Model

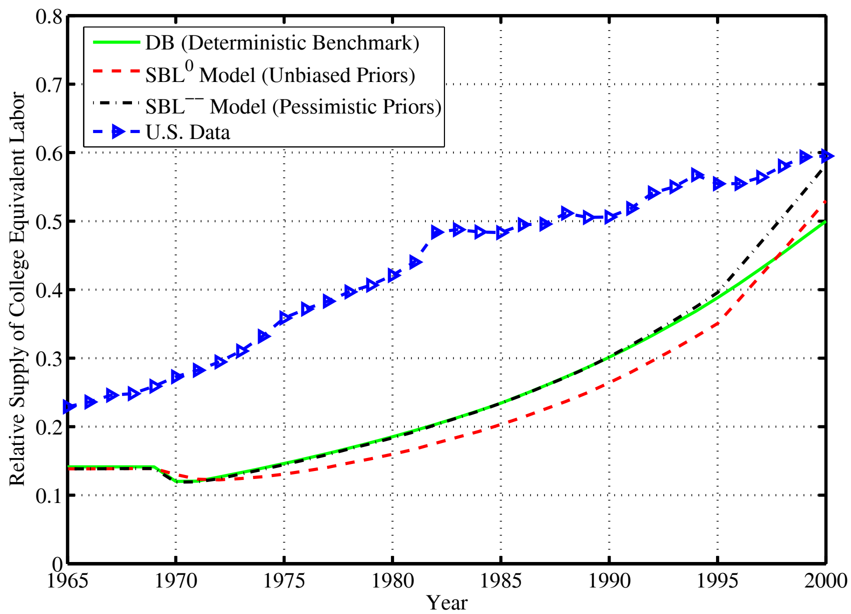


before SBTC, which is consistent with the data. After SBTC, the initial decline and the strong rise in college premium is apparent among younger workers, but there is no fall among more experienced workers and the rise is slightly smaller as well. As a result, the gap between the older and younger workers widens initially (from 0.36 to 0.45) and then narrows to 0.22 in 1990. This result is largely due to the fact that young individuals—who face a longer planning horizon, and hence have a larger marginal benefit from investing—respond to SBTC more than older individuals. In contrast, the increase in the college premium among older workers is mainly driven by price effect: the 30-45 experience group in 1995 had 10 to 25 years of experience in 1975, so their investment response is minimal. Therefore, in the short-run the investment effect (and in the long-run, the quantity effect) has a smaller negative (positive) impact on the college premium among older workers.

The increase in the college premium among older workers in the model is more pronounced than in the U.S. data. One factor that could explain this discrepancy is the fact that our model assumes that all SBTC has been "disembodied." As a result, older workers who accumulated human capital before 1970 stood to gain as much as newer cohorts per unit of human capital. Instead, if some part of SBTC was embodied in new vintages of machines and new vintages of human capital associated with these machines, then older cohorts would have benefitted less from SBTC, resulting in a smaller increase in the college premium.[17] Although, such an alternative modeling would also have different

[17] See Greenwood and Yorukoglu (1997), Caselli (1999) and Violante (2002) for examples of models where improve-

Figure 7: The Relative Supply of College Equivalent Labor, 1965—2000: Model versus U.S. Data



implications for the questions examined so far, the idea of analyzing human capital decision in an environment with *embodied* SBTC seems interesting and deserves further examination.

3.2.4 The Rise in the Relative Supply of College Labor

Another prominent trend during this period has been the significant rise in educational attainment and the consequent increase in the relative supply of labor hours worked by college-educated individuals. This trend can be seen in figure 7 where the dashed line (marked with triangles) plots 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. Several studies have emphasized exogenous driving forces behind this rising supply, and viewed the evolution of the college premium as resulting from the interaction of this supply growth together with the demand change due to SBTC (c.f., Freeman (1976), Katz and Murphy (1992), Acemoglu (1998)).

In the present model, instead, college enrollment is determined endogenously, and in particular, it responds to the change in the price of skills resulting from SBTC. The three closely spaced lines

ments in technology comes in the form of new vintages of machines that are more productive.

in figure 7 show the relative supply of college-equivalent labor implied by the model.[18] The first point to observe is that all three versions of the model understate the level of relative supply prior to SBTC. Perhaps this does not come as a surprise, since no attempt was made to match the relative supply in the calibration. However, this supply grows significantly during SBTC, because human capital investment (and hence educational attainment) is increasing for every successive cohort, and over time these younger cohorts are replacing older ones with lower educational attainment. As a result, relative supply increases by a total of 0.36 in the DB model, which compares well with the rise of 0.35 in the data. This similarity is surprising given that in the model educational attainment is modeled merely as a by-product—depending on whether investment exceeds a certain threshold or not—and many potentially important aspects of education have been left out, such as tuition costs, tuition subsidies, changes in the quality of education, changes in cohort sizes, etc. This analysis suggests that changes in the price of skills might have played a more important role than these factors in determining the overall rise in educational attainment. In the stochastic models (SBL^0 and SBL^-) the rise in enrollment is very similar to the deterministic benchmark until 1995, but grows faster after that date.[19]

One aspect of the data that is not explained well by the model is the slowdown in the growth of supply starting in the 1980's. Part of the fast rise in supply before the 80's has been attributed to the large sizes of baby-boom cohorts (together with the secular trend in college enrollment over time) as well as to the Vietnam Era draft laws which made college enrollment more attractive (c.f., Acemoglu (2002)). These features are not present in the model which could explain part of the discrepancy with the data.[20]

3.2.5 Within-group Inequality

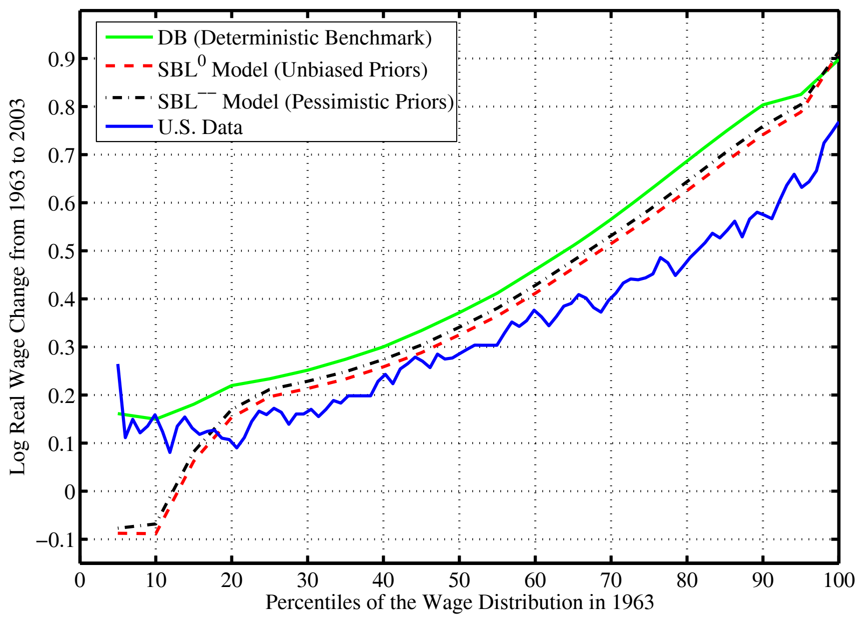
The analysis so far has focused on the evolution of some key moments of the wage distribution. However, a distribution typically contains much more information than what can be summarized by a few moments, and it is possible for a model to be consistent with some summary statistics,

[18] Relative supply is defined as the ratio of working individuals ($i < \chi$) who have completed more than two years of full-time investment ($i = 1$), to those who have had less investment. This is analogous to the definition adopted by Autor et al (2005) when constructing the empirical counterpart.

[19] This is because individuals do not foresee the slowdown in the growth of the price of human capital after that date, and continue to invest at the same rate as before (whereas in the DB model they know immediately when θ_H stops growing). This is an unintended consequence of our attempt to simplify the learning problem, and does not seem to be very interesting to deserve further comment.

[20] There are some other features of education and human capital accumulation that have been abstracted away in the model which could also be responsible for a slower growth in college enrollment in the data. For example, in reality learning requires more than time: it requires school buildings, equipment, teachers, and so on. Many of these inputs could have inelastic supply in the short-run and even in the intermediate-run. The rise in college tuitions and the relative wages of educators in the last several decades could be indicative of demand pressures on inputs whose supply may have limited elasticity.

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

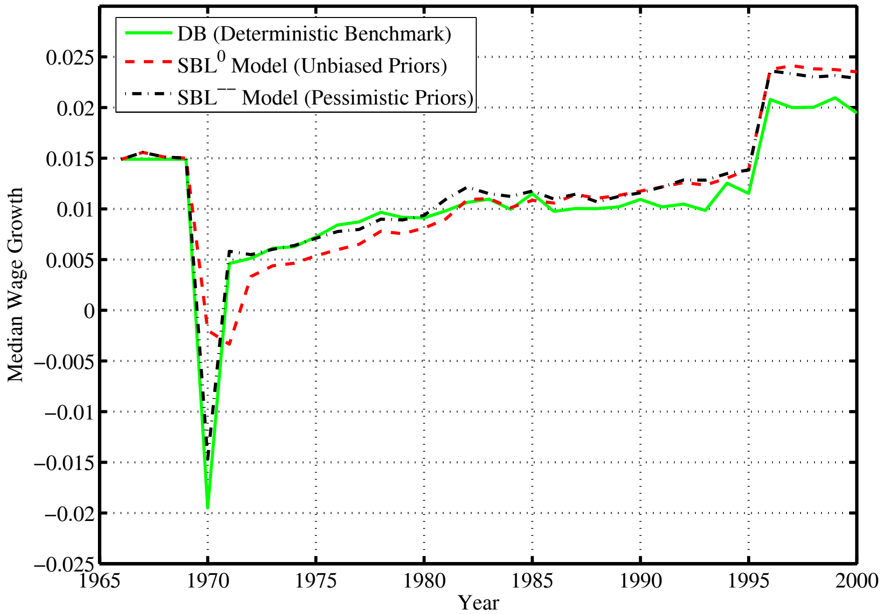


but generate patterns inconsistent with the data at a more disaggregated level. Juhn, Murphy and Pierce (1993) have documented a striking empirical regularity at a very disaggregated level which presents such a challenge. In figure 8 we report the same finding using our data set which covers a longer 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.[21] 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 counterparts are also plotted in figure 8. They show the same general pattern of widening inequality that is spread quite evenly across the wage distribution as observed in the data (with the exception of the lower tail discussed below). The mechanism behind this result should be clear from earlier discussions. Wage inequality arises entirely from differences in human capital

[21] Juhn, 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 corresponding graphs from our model that qualitatively look similar to the data. We do not discuss them for brevity; they are available upon request.

Figure 9: The Growth Rate of Median Wages in the Model, 1965-2000



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. The existence of this same pattern in the data suggests that this mechanism could be an important channel behind the rise in within-group inequality.

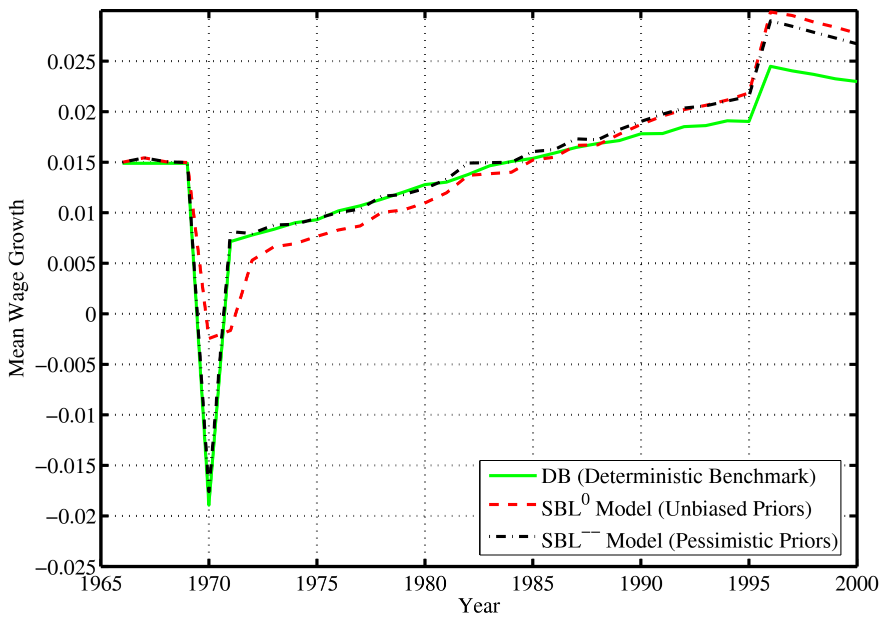
The only noticeable difference between the different versions of the model is that while the deterministic DB model matches the wage growth in the lower 20 percentiles quite well, the models with Bayesian learning (SBL^0 and SBL^-) underestimate it quite significantly. Finally, one discrepancy between all three models and the data is the higher *average* wage growth in the model, which is discussed next.

3.3 Evolution of Average Wages

3.3.1 Productivity Slowdown and Stagnation of Median Wages

Macroeconomists and labor economists have documented two closely related trends during this period: the slowdown in labor productivity and the stagnation of median wage growth, which both started with a sharp fall in 1973 and persisted until about 1995. For example, Juhn, Murphy and

Figure 10: The Growth Rate of Labor Productivity in the Model, 1965—2000



Pierce (1993) report that the median real wage has increased by 2.2 percent per year between 1963 and 1973, but actually *fell* by about 0.3 percent per year between 1973 and 1989. Similarly, labor productivity (measured as the non-farm business output per hour) has grown by 2.6 percent per year from 1955 to 1973, but only by 1.45 percent per year from 1973 to 1995.[22]

Figures 9 and 10 plot, respectively, the growth rates of median wages and labor productivity implied by the model.[23] First, both series fall sharply immediately after SBTC starts in 1970. Hence, the model is able to generate the sharp initial slowdown, although this happens three years earlier than in the data.

After the initial fall, the median wage continues to stagnate: in the DB model, it grows at 0.46 percent per year from 1970 to 1979, and averages 0.81 percent overall until 1995, representing a significant slowdown compared to the 1.5 percent growth during the period before 1970. As can be seen in the graph, the corresponding figures from the SBL^0 and SBL^- models are very similar to the DB model. The only noticeable difference is in the SBL^- model, where the initial crash in median wage growth is more modest, but this is compensated by an overall slower growth during the 1970's. As a result, the stagnation of median wage growth in the 1970's is 0.45 percent

[22] Authors' calculation from Bureau of Labor Statistics data.

[23] Notice that since there is no capital in the model total output equals total wages, implying that labor productivity (output per hour) equals the mean wage rate in the economy.

and 0.36 percent in the SBL^0 and SBL^- models respectively. Similarly, in the DB model, labor productivity grows by only 0.67 percent per year during the 1970's, but recovers faster and averages 1.26 percent per year until 1995. The other two models behave similarly again. Overall then, while the magnitude of slowdown is somewhat smaller than in the data, the model correctly predicts the qualitative aspects of this evolution, including the sharp initial fall, the sluggish nature of the subsequent recovery, and the fact that the slowdown was larger for median wages than it was for labor productivity.

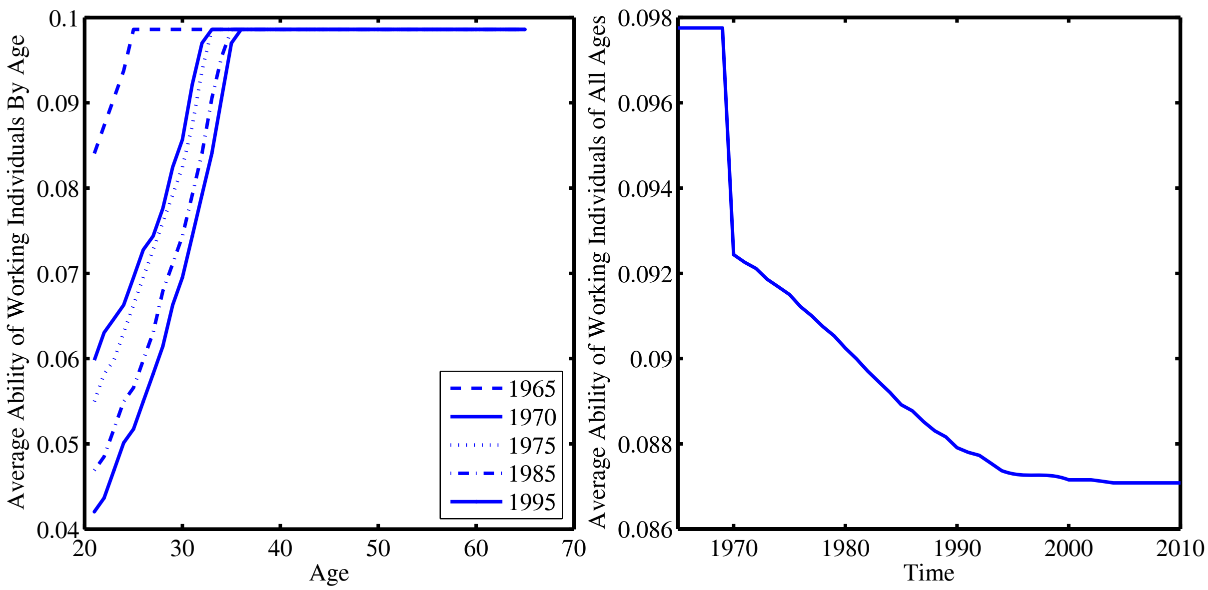
There are two related, yet different, channels that play an important role. To see this, recall that the increase in investment after SBTC can take one of two forms. First, it will increase both the fraction of individuals who invest full time (i.e., enroll in college) and lengthen the duration for those already planning to go to college. Since this change takes place at the upper tail of the ability distribution, the average ability of individuals who remain in the labor market continually falls during SBTC, as can be seen in the right panel of figure 11. Because, individuals with lower ability also have low human capital on average, this "selection effect" reduces average wages and productivity after SBTC. The magnitude of this selection effect seems empirically plausible as evidenced by the fact that the model roughly matches the total growth in college enrollment rate during this period (figure 7).

Second, those who remain in the labor market also respond to SBTC by increasing their on-the-job investment. This is shown in the right panel of figure 12. The fraction of time invested before SBTC is 7.2 percent (or 2.9 hours in a 40-hour work week) and increases to reach 13.1 percent in 1995 (or 5.2 hours a week). Neither the initial investment level, nor the increase during SBTC appears implausibly large to us, 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 is that the investment response is concentrated among younger individuals (left panel of figure 12) and thus the change in *average* investment is small. Another reason is that on-the-job investment is bounded from above by $\chi = 0.50$. Nevertheless, this "on-the-job investment effect" works in the same direction as the "selection effect" to further reduce wage growth after SBTC.

An important result that follows from this discussion is that the choice of χ does not play a critical role here. A higher value of χ would allow for a larger increase in "on-the-job investment," but this would be offset by a smaller change in the ability composition because fewer individuals would now leave the labor market for full-time education. For example, setting $\chi = 0.80$ results in a median wage growth of 0.77 percent (compared to 0.82 above), and a productivity growth of 1.23 percent (compared to 1.24 above).

To sum up, during this period the labor market is dominated by individuals who have lower ability than before, but who also invest more than before, resulting in slow wage and productivity

Figure 11: Evolution of the Average Ability of *Working* Individuals Over Time

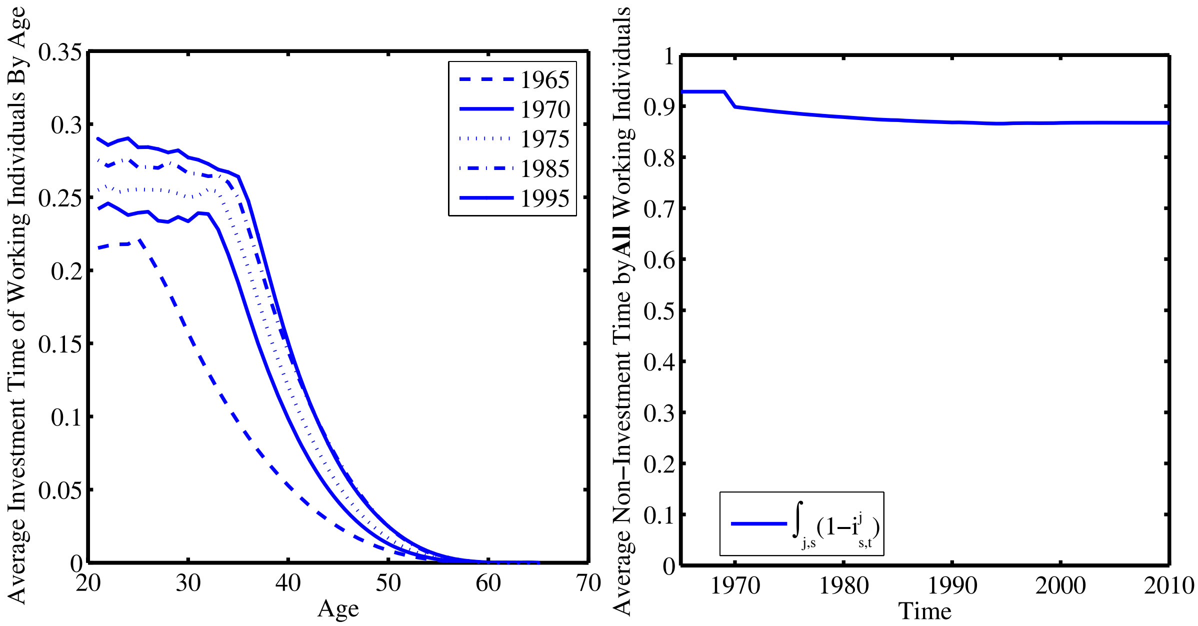


growth. Over time, the increase in the total human capital stock due to both types of investment begins to dominate, resulting in a recovery in both the median wage and labor productivity. The model also predicts that the stagnation results, loosely speaking, from the response of *total* human capital investment to SBTC. What fraction of this investment happens on-the-job or at school has a small effect on the broad picture.

3.3.2 Cross-sectional Wage Profiles by Education and Experience

Another set of well-documented trends during this period concern cross-sectional wage profiles. To discuss these facts, in Table 2 we reproduce the relevant figures from Katz and Murphy (1992, table 1). The table reports the average wage growth for different education-experience groups over time. Perhaps the most striking fact—noted by several authors—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 plummet by 19.8 percent while older workers see only a small fall of 2.8 percent. As a result, the average 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 average cross-sectional wage profile

Figure 12: Evolution of the Average Investment of *Working* Individuals Over Time



flattens for this group.[24] See also Bound and Johnson (1992, Figure 1).

We construct the model counterparts (from the *DB* model) of the same statistics with one difference. As discussed in the previous section 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, not only during the 1980's but also going back to the 1970's, rather well. For example, among high-school graduates there is little difference in wage growth by experience levels during the 1970's, whereas for college graduates there is a larger fall for younger individuals than for older ones. More importantly, the model is also consistent with the wage changes of all four education-experience groups from 1979 to 1987 noted above. As a result, the average cross-sectional wage profile steepens for individuals with low education and flattens for those with high education during this period.

There are three effects that should be taken into account to understand the wage changes of

[24] Clearly, these facts are closely related to the evolution of the college premium within age groups discussed above. However, notice that the college premium is only informative about the *relative* wages of these two groups, whereas the current facts relate separately to the evolution of the *levels* of each group's wages.

Table 2: Real Wage Changes By Education and Experience Groups, 1971-1987

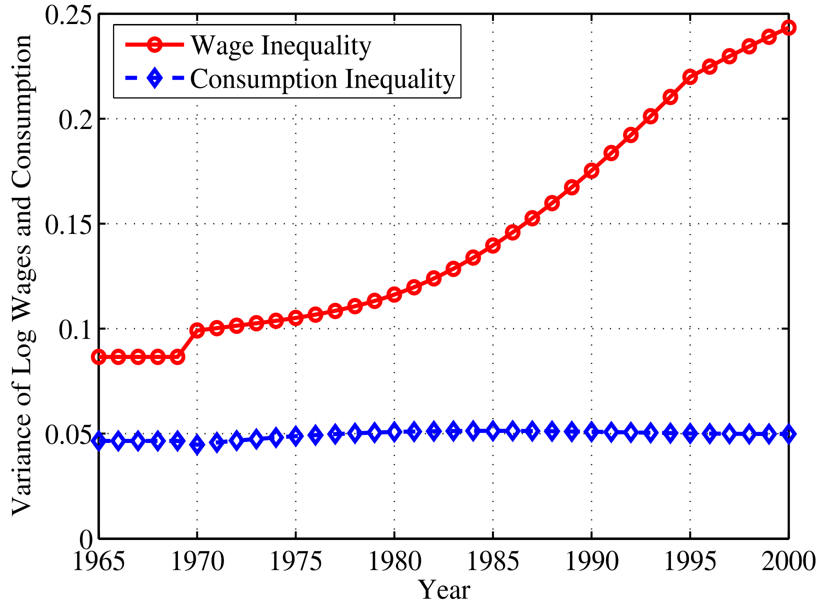
Group			Change in Log Average Real Wage (multiplied by 100)	
Education	Experience	Sample	1971-79	1979-87
12	Low	Data	**0.8**	**−19.8**
12	Low	Model	−1.9	−8.7
12	High	Data	**3.2**	**−2.8**
12	High	Model	−1.2	−3.7
16+	Low	Data	**−11.3**	**10.8**
16+	Low	Model	−7.8	13.3
16+	High	Data	**−4.0**	**1.8**
16+	High	Model	3.8	2.6

Notes: The empirical statistics reported are taken from Katz and Murphy (1992, Table 1). The model counterpart is from the DB model. 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.

high-school graduates. *First*, young high school graduates also respond to SBTC—even if it is not to the same extent as high-ability individuals—by increasing their investment. Because the majority of these individuals remain below the ability threshold for college enrollment, most of this investment takes place on-the-job, which reduces their observed 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 channel are a problem for *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 offset their loss on their raw labor endowment. A combination of these three factors, which work in opposite directions for the young and old, explain why the former group experienced a large wage loss while the latter saw no significant change during the 1980's. 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 (1980's).

The mechanism for the behavior of the wages of college graduates is similar, but the existence of an upper bound on on-the-job training also plays a role. This is because, after SBTC high-ability individuals who want to increase their investment significantly have to stay in college longer due

Figure 13: The Evolution of Wage and Consumption Inequality in the Model: 1965—2000.



to the upper limit on investment on the job. 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.

3.4 Evolution of Lifetime Income (Consumption) Inequality

A somewhat surprising empirical finding from this period is that the rise in consumption inequality has been muted compared to the rise in wage inequality. Although there remains some disagreement about the exact magnitude of the rise in consumption inequality (mainly due to data problems), several authors report findings broadly supporting this conclusion (see for example Krueger and Perri (2005) 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 analysis of consumption inequality (such as incomplete markets, retirement savings, demographic changes, etc.). But the model still addresses a simple but fundamental question: Has the substantial rise in cross-sectional wage inequality during this period resulted in a parallel rise in *life-time* income inequality? Figure 13 plots the evolution of life-time income (which equals consumption in the model) inequality in the DB model, which shows a very small increase of 0.2 log points during SBTC. [The results from the SBL^0 and SBL^- models are to be added].

At first blush, it seems quite surprising that wage inequality could rise in such a systematic fashion without a significant change in life-time incomes. The mechanism can be anticipated from the earlier discussion of figure 14. First, because wage inequality increases by an increased dispersion in *growth rates*, life-time inequality—which is the variance calculated after averaging wages over the life-cycle—increases by less. This would not be the case if the increase in dispersion was in the levels of the wage profiles, in which case, consumption inequality would increase one for one with wage inequality. Second, and furthermore, the higher wages of high-ability individuals later in life come at a large cost in the form of high investment and low wages early on, driving down the lifetime gain from human capital investment. As shown in Kuruscu (2005), for a range of plausible parameter values similar to those used here, the gain in lifetime income due to human capital investment is surprisingly small—about 1 percent. This is because the foregone earnings during the high investment early in life is close to the future discounted benefits. Therefore, this model offers a mechanism that is consistent with a large increase in wage inequality but a small change in life-time inequality.

It is possible to generate a larger increase in lifetime inequality in this model under alternative parameterizations. However, for parameter values broadly consistent with facts about the wage distribution, the largest increase we obtained was less than 5 log points. Introducing depreciation in human capital also works to generate a larger increase in consumption inequality. A fuller investigation of this model for consumption facts is left for future work. [Results from SBL^0 and SBL^- will be added].

4 Extensions and Robustness

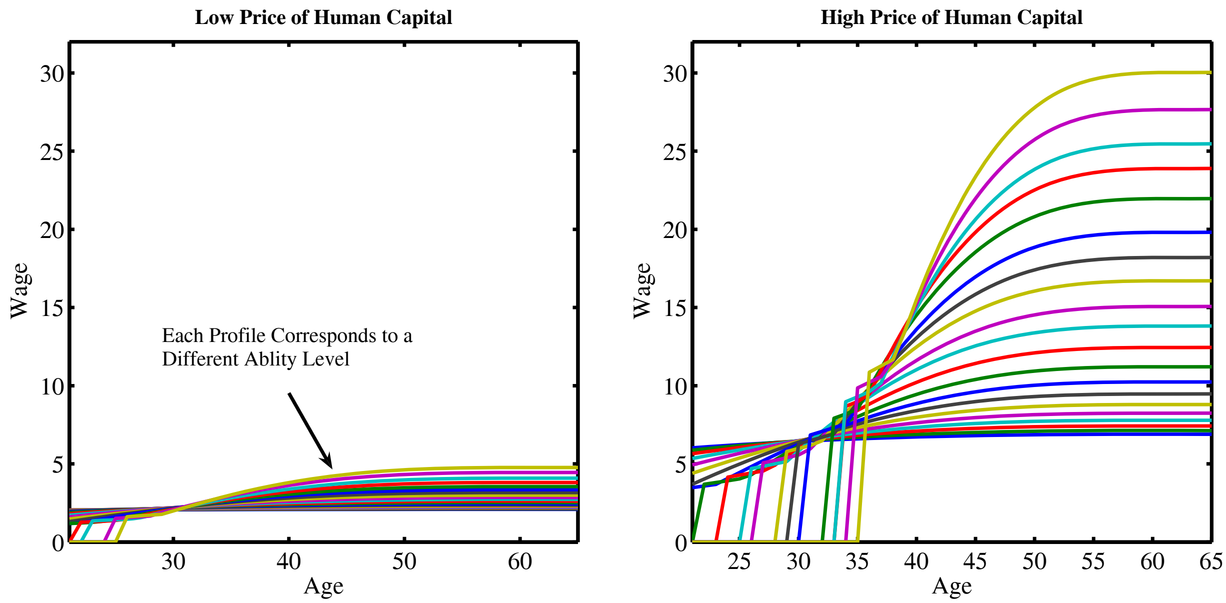
4.1 Allowing for Imperfect Substitution

[TO BE ADDED]

4.1.1 Calibrating the Curvature of Production Function

We are not aware of any estimate of ρ in the literature using the production function in (3) that would guide our calibration. One difficulty with directly estimating ρ from data is that our production function features inputs such as human capital and investment time that are very difficult to measure directly in the data. However, there is a large literature that has estimated a

Figure 14: The Effect of the Returns to Human Capital on Life-Cycle Wage Profiles



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 (c.f., Katz and Murphy (1992). See Hamermesh (1993) for a survey). This elasticity, denote it with ϕ, is obtained by running the following regression:

$$\log \omega_t^* = a_0 + a_1 t - \frac{1}{\phi} \log \left(N_{c,t}/N_{nc,t} \right) + error, \tag{11}$$

where N_c/N_{nc} is the labor supply of college-educated workers relative to non-college workers, using the notation developed above. We take $\widehat{\phi} = 2$ as our empirical benchmark value, consistent with Hamermesh's (1993) survey of a large number of studies that estimate this parameter.[25]

Notice that all the variables that appear in this regression can be generated from our model as well, which suggests one way to calibrate ρ. Essentially, we can choose ρ such that when we run the same regression above using simulated data from our model, we obtain the same estimate of ϕ as in these earlier studies. This approach however presents its own challenges. The main difficulty is that when $\rho < 1$ and there are aggregate shocks (which would be required to run the regression above), the equilibrium prices of raw labor and human capital will depend on H and L, which

[25] Similarly, Autor, Katz and Kearney (2005a) use the same data as in the present paper and estimate values of $\widehat{\phi}$ that range between 1.66 to 2.05 (see Table 2, columns 2 and 3 of their paper). However, it should be kept in mind that these studies pool data on males and females under the implicit assumption that the two are perfect substitutes in the CES procudtion function. Instead here we only use data on male wages and relative supply.

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

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

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

in turn depend on the distributions of human capital and investment time in the population (see equations (2) and (4)). One would then have to track the evolutions of these distributions over time, as they become state variables of the model (and use an algorithm such as the one developed by Krusell and Smith (1998)). Given this additional complexity, and given that in the previous section $\rho = 1$ delivered quite plausible implications, it is useful to first examine how the elasticity $\widehat{\phi}$ implied by this choice compares to empirical estimates.

Table 3 reports the results when $\rho = 1$ in our model. To obtain these results we proceeded as follows. We solved the stochastic version of the model (as in the SBL models above), but without SBTC, and therefore, also without heterogeneity in beliefs (that is we set $\kappa = 0$ throughout in equation (6), and $\nu \equiv 0$). We chose the standard deviation of skill-bias shocks, σ_ε^2, as in the baseline case above (to be consistent with the volatility of the college premium in the U.S. data from 1963 to 2003). Then we simulated data for 30 years and estimated the regression in (11), 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 $\widehat{\phi} = 2.05$ with an R^2 of 0.94, in the closest specification to ours (see column 3 of table 2). It seems surprising that a choice of $\rho = 1$, which implies perfect substitution between H and L in our model (in other words, *infinite* substitution elasticity!) not only generates a finite elasticity between college and non-college workers, but also a relatively small value similar to that observed in the data.

To understand these results, notice that in our model $\rho = 1$ implies perfect substitution between different workers in a very specific and limited sense: two workers will be perfectly substitutable only in an environment where individuals' (unobserved) human capital and investment are constant. To see this, consider two individuals who differ in their age and ability levels: (s, j) and (s', j'). To keep a constant output level, one would need to substitute Υ workers of type (s', j') for 1 worker of type (s, j) where:

$$\Upsilon\left((s,j),(s',j')\right) = \frac{\left[\theta_L l + \theta_H h_{s'}^{j'}\right]\left(1 - i_{s'}^{j'}\right)}{\left[\theta_L l + \theta_H h_s^{j}\right]\left(1 - i_s^{j}\right)},$$

is the marginal rate of substitution in production between these two types of workers. Note that Υ will only be constant when $h_s^j, h_{s'}^{j'}, i_s^j$ and $i_{s'}^{j'}$ are all constant. But this is not going to be the case unless the economy is in a complete steady state. When skill prices move over time, either due to SBTC or due to stochastic shocks to skill prices, the response of different individuals' investment will be systematically different depending on their age and ability, in turn creating systematic changes in Υ.

To understand the results of the regression above, consider the response of the 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 investment. At the extensive margin, more individuals enroll in college which 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. As explained in the previous section, this results in a fall in the college premium. Therefore, following a positive ε shock, the supply of college workers rises while the college premium falls. This negative correlation makes it appear as if a 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 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 the change in the supply of college labor were to happen for completely exogenous reasons in our model (and the workers to be added were selected randomly from among college workers, or among non-college workers) then the estimate of $(1/\phi)$ would be zero, revealing an infinite supply elasticity. [TO BE COMPLETED]

4.2 What Happens if Investment does not Respond to SBTC?

In the previous sections, we have noted that the response of investment behavior to SBTC was critical for the results of this paper. Instead, earlier studies have typically placed a lot of emphasis on the role of rising skill prices for the evolution of wage inequality, and changes in quantities have often been viewed as secondary (see especially Juhn et al (1993)). Thus, to shed further light on the role of prices and quantities, we now examine the implications of our model under the assumption that human capital investment does not change after SBTC (which we refer to as the "fixed-investment model"). Specifically, we assume that while θ_H/θ_L starts to increase after 1970

Figure 15: Evolution of Key Variables when Investment does Not Respond to SBTC

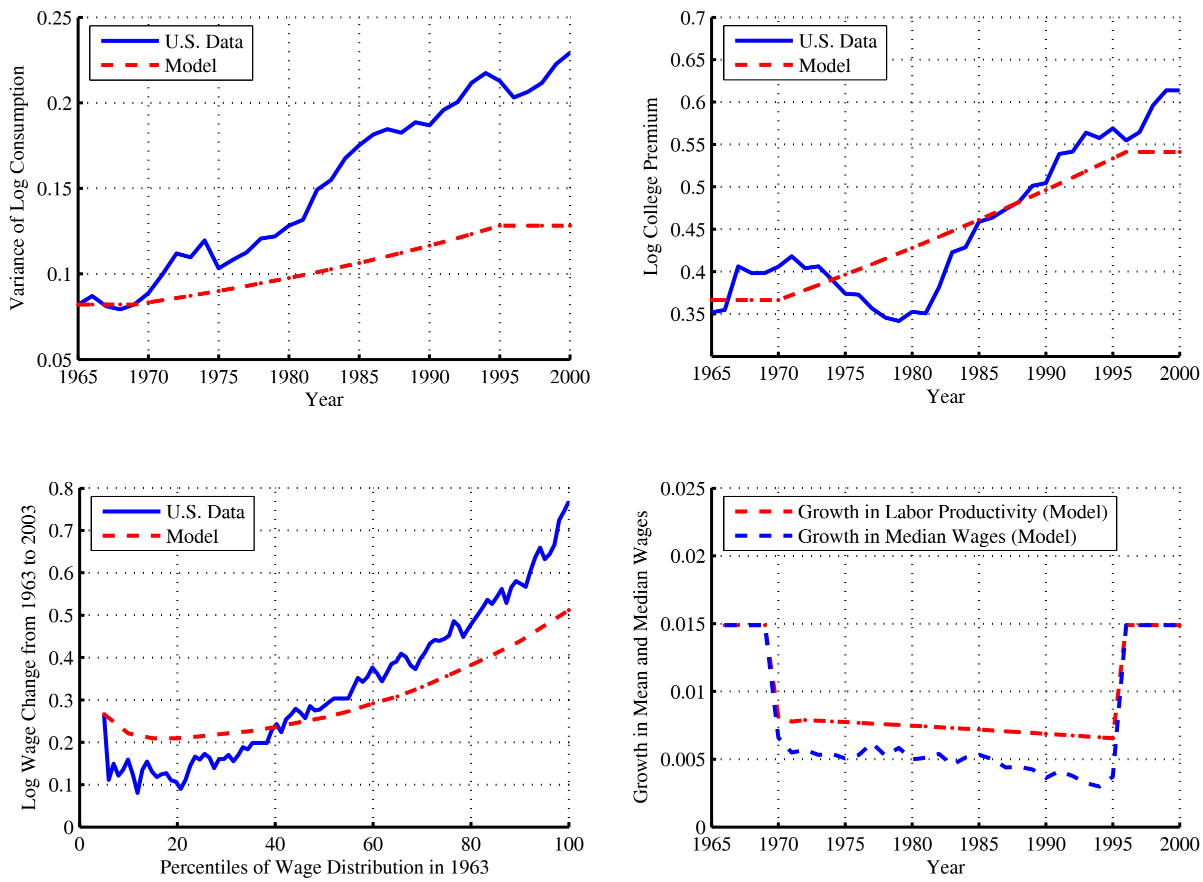


Table 4: EVOLUTION OF INEQUALITY WITHOUT INVESTMENT RESPONSE TO SBTC

	Benchmark Model		Fixed Investment Model	
Year	$Var(log(w))$	$Var(log(c))$	$Var(log(w))$	$Var(log(c))$
69	0.0821	0.0440	0.0821	0.0440
95	0.2087	0.0458	0.1282	0.0451
Difference	0.1266	0.0018	0.0461	0.0011

(as in the baseline experiment), all individuals continue to invest at the same rate as before SBTC. In this case, the distribution of human capital across individuals remains unchanged and the only effect of SBTC is through the price of human capital.

We first compare the rise in wage inequality and consumption inequality in the fixed-investment model to the benchmark case. As shown in Table 4, the rise in wage inequality between 1969 and 1995 is about 4.6 log points in the fixed-investment model, implying that roughly 36 percent (4.61/12.66) of the increase in wage inequality is due to price effect, while the remaining 64 percent results from the response of investment. This picture is reversed when it comes to consumption inequality, which rises in the fixed-investment model by roughly 61 percent of the baseline model. Therefore, a smaller fraction (39 percent) of the rise in consumption inequality is attributable to investment behavior and a larger fraction results from changes in skill prices (Of course the rise in consumption inequality is very small in both cases). This finding is consistent with the intuition discussed in the previous section that human capital investment does not increase lifetime income substantially because those who experience higher wages later in life are the ones who invest more and accept lower wages early in life.

Figure 15 displays several statistics that we are interested in. As can be seen in the figure, about one-third of the increase in wage inequality is due to changes in prices. Second, although the magnitude of the total rise in the college premium is roughly similar to the benchmark case, there is no fall in the college premium during the 70's. This confirms our assertion that the response of investment to SBTC is critical for the initial decline in the college premium. Third, wage changes at different percentiles of the distribution shows that investment response is especially important for the growth of wages at the upper end of the distribution. Fourth, the slowdown in labor productivity and the stagnation of median wage growth are more pronounced when investment does not respond to SBTC, which is intuitive. To sum up, if investment did not respond to SBTC, the increase in wage inequality would be much smaller, the college premium would increase monotonically and the slowdown in productivity would be further protracted.[26]

[26] Of course it is possible to recalibrate the rise in θ_H/θ_L so that the model matches the increase in wage inequality from 1970 to 1995. However, this calibration has several counterfactual implications, such as an increase in the college

4.3 Positive Depreciation of Human Capital

[To be written]

5 Conclusion

[To be written]

References

[1] Acemoglu, Daron (1998): "Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality," *Quarterly Journal of Economics,* Vol. 113, pp. 1055-1089

[2] Acemoglu, Daron (2002): "Technical Change, Inequality, and Labor Market", *Journal of Economic Literature*, March 2002, pp. 7-72.

[3] Aghion, Philippe, Peter Howitt, and Gianluca Violante (2002): "General Purpose Technology and Wage Inequality," *Journal of Economic Growth,* Vol. 7(4), pp. 315-345.

[4] Attanasio, Orazio, Erich Battistin, and Hidehiko Ichimura (2004): "What really happened to consumption inequality in the US?" IFS Working paper.

[5] Autor, David, Lawrence F. Katz, Alan B. Krueger (1998): "Computing Inequality: Have Computers Changed the Labor Market?" *Quarterly Journal of Economics.* 113, pp. 1169-1214.

[6] 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.

[7] Autor David H., Lawrence F. Katz, and Melissa S. Kearney (2005a): "Trends in U.S. Wage Inequality: Re-Assessing the Revisionists," MIT working paper.

[8] Autor David H., Lawrence F. Katz, and Melissa S. Kearney (2005b): "Rising Wage Inequality: The Role of Composition and Prices," MIT working paper.

[9] Baker, Michael (1997): "Growth-rate Heterogeneity and the Covariance Structure of Life-Cycle Earnings," *Journal of Labor Economics,* Vol. 15, pp. 338-375.

[10] Beaudry, Paul and David A. Green (2003): "Wages and Employment in the United States and Germany: What Explains the Differences?" *American Economic Review*, Vol. 93, No. 3.

premium that is unrealistically large, productivity growth that is negative from 1970 to 1995, among others.

[11] Becker, Gary S. (1965): "Human Capital : A Theoretical and Empirical Analysis, with Special Reference to Education," reprinted 1994, University of Chicago Press.

[12] Ben Porath, Yoram (1967): "The Production of Human Capital and the Life Cycle of Earnings." *Journal of Political Economy*, 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 Industries: Evidence from the Annual Survey of Manufactures." *Quarterly Journal of Economics,* Vol 109, pp. 367-365.

[14] Blanchard, Olivier (1985): "Deficits, Debt and Finite Horizons," *Journal of Political Economy,* Vol 93, p. 223-247.

[15] Blundell, Richard and Ian Preston (1998): "Consumption Inequality and Income Uncertainty," *Quarterly Journal of Economics*, Vol. CXIII, Issue 2, 603-640.

[16] Bound, John and George Johnson (1992): "Changes in the Structure of Wages in the 1980's: An Evaluation of Alternative Explanations," *American Economic Review,* Vol 82, pp. 371-92.

[17] Card, David, and John Dinardo (2002): "Skill Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles," *Journal of Labor Economics,* Vol 20, pp. 733-83.

[18] 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*, 116, pp. 705-46.

[19] Caselli, Francesco (1999): "Technological Revolutions," *American Economic Review,* pp. 78–102.

[20] Davis, Steve J., Felix Kubler and Paul Willen (2003): "Borrowing Costs and the Demand for Equity Over the Life-cycle," Working Paper, University of Chicago.

[21] Freeman, Richard (1976): "The Overeducated American," Academic Press.

[22] Galor, Oded and Omer Moav (2002): "Ability Biased Technological Transition, Wage Inequality and Economic Growth," *Quarterly Journal of Economics,* Vol 115, pp. 469-497.

[23] Galor, Oded, and Daniel Tsiddon (1998): "Technological Progress, Mobility and Economic Growth," *American Economic Review.* Vol 87, pp. 363-382.

[24] Gould Eric, Omar Moav, and Bruce Weinberg (2001): "Precautionary Demand for Education, Inequality, and Technological Progress," *Journal of Economic Growth*, 6, pp. 285-315.

[25] Gourinchas Pierre-Olivier, and Jonathan A. Parker (2002): "Consumption Over the Life Cycle," *Econometrica,* Vol. 70, pp. 47-89.

[26] Greenwood, Jeremy, and Mehmet Yorukoglu (1997): "1974," *Carnegie Rochester Conference on Public Policy,* Vol 46, pp 49-94.

[27] Guvenen, Fatih (2005a): "An Empirical Investigation of Labor Income Processes," Working paper, University of Texas at Austin.

[28] Guvenen, Fatih (2005b): "Learning Your Earning: Are Labor Income Shocks Really Very Persistent?, Working Paper, University of Texas at Austin.

[29] Guvenen, Fatih, and Burhanettin Kuruscu (2006): "Ben-Porath Meets Skill-Biased Technical Change: A Theoretical Analysis of Rising Inequality," Working Paper, University of Texas at Austin.

[30] Heathcote, Jonathan, Kjetil Storesletten, Gianluca Violante (2004): "The Macroeconomic Implications of Rising Wage Inequality in the United States," Working Paper, NYU.

[31] Heckman, James J. (1976): "A Life-cycle Model of Earnings, Learning, and Consumption," *Journal of Political Economy* 84(4), 11-44.

[32] Heckman, James J.; Lochner, Lance J. and Taber, Christopher R. (1998): "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogenous Agents." *Review of Economic Dynamics*, 1(1), pp. 1-58.

[33] Huggett, Mark, Gustavo Ventura, and Amir Yaron (2005): "Sources of Life-Cycle Inequality," Working Paper, Wharton School, University of Pennsylvania.

[34] Juhn, Chinchui, Kevin M. Murphy, and Brooks Pierce (1993): "Wage Inequality and the Rise in Returns to Skill," *Journal of Political Economy*, pp. 410-442.

[35] Kambourov, Gueorgui and Iourii Manovskii (2005): "Accounting for the Changing Life-Cycle Profile of Earnings," Working Paper, University of Pennsylvania.

[36] 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, Volume 3A, Amsterdam: Elsevier-North Holland, 1463 –1555.

[37] Katz, Lawrence F., and Kevin Murphy (1992): "Changes in Relative Wages: Supply and Demand Factors," *Quarterly Journal of Economics,* pp. 35-78.

[38] Krueger, Dirk and Fabrizio Perri (2006): "Does Income Inequality Lead to Consumption Inequality? Some Empirical Findings and a Theoretical Explanation," *Review of Economic Studies*, Vol. 73(1), 163-193.

[39] Krusell, Per, Lee Ohanian, Victor Rios-Rull, and Gianluca Violante (2000), "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis" *Econometrica*, Vol. 68, pp. 1029-1054.

[40] Kuruscu, Burhanettin (2005): "Training and Lifetime Income," *American Economic Review*, Vol. 96, No 3, pp. 832-846.

[41] Lillard, Lee A., and Yoram A Weiss (1979): "Components of Variation in Panel Earnings Data: American Scientists, 1960-70," *Econometrica*, Vol. 47, pp. 437-454.

[42] Meghir, Costas, and Luigi Pistaferri (2004): "Income Variance Dynamics and Heterogeneity," *Econometrica*, Vol. 72, pp. 1-32.

[43] Machin, Stephen and John Van Reenen (1998): "Technology and Changes in Skill Structure: Evidence from Seven OECD Countries." *Quarterly Journal of Economics.* 113:4, pp. 1215-44.

[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, pp. 285-326.

[46] Violante Giovanni L. (2002): "Technological Acceleration, Skill Transferability, and the Rise in Residual Inequality," *Quarterly Journal of Economics,* Vol. 117, pp 297-338.