Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents*

James J. Heckman[†] and Lance Lochner

Department of Economics, University of Chicago, Chicago, Illinois 60637

and

Christopher Taber

Department of Economics and Institute for Policy Research, Northwestern University, Evanston, Illinois 60208

This paper develops and estimates an overlapping generations general equilibrium model of labor earnings, skill formation, and physical capital accumulation with heterogenous human capital. The model analyzes both schooling choices and post-school on-the-job investment in skills in a framework in which different schooling levels index different skills. A key insight in the model is that accounting for the distinction between skill prices and measured wages is important for

analyzing the changing wage structure, as they sometimes move in different directions. New methods for estimating the demand for unobserved human capital and for determining the substitution relationships between skills and capital in aggregate technology are developed and applied. We estimate skill-specific human capital accumulation equations that are consistent with the general equilibrium predictions of the model. Using our estimates, we find that a model of skill-biased technical change with a trend estimated from our aggregate technology is consistent with the central feature of rising wage inequality measured by the college–high school wage differential and by the standard deviation of log earnings over the past 15 years. Immigration of low-skill workers contributes little to rising wage inequality. When the model is extended to account for the enlarged cohorts of the Baby Boom, we find that the same parameter estimates of the supply functions for human capital that are used to explain the wage history of the last 15 years also explain the last 35 years of wage inequality as documented by Katz and Murphy (L. Katz and K. Murphy, *Quart. J. Econ.* **107** (1992), 35–78). *Journal of Economic Literature* Classification Numbers: J24, J31, D58, D33.   © 1998 Academic Press

Wage inequality has increased substantially in the United States since the early 1960s. Workers with low skills and little education have experienced large declines in their earnings, both absolutely and relative to more skilled workers. Only recently have economists begun to develop models that explain the rise in wage inequality, focusing on the college–high school wage differential. The primary causes for the recent increase in overall wage inequality are still unknown.

This paper develops a heterogeneous-agent dynamic general equilibrium model of labor earnings, estimates the model using micro data, and uses the estimated model to explore the empirical plausibility of alternative explanations for the recent rise in wage inequality. Our model has several sources of heterogeneity among its agents: (1) Persons differ in initial ability levels and this ability affects both earnings levels and personal investment behavior. (2) Skills are heterogeneous. Different schooling levels correspond to different skills. The post-school skills acquired at one schooling level are not perfect substitutes for the post-school skills acquired at another schooling level, but skills are perfect substitutes across age groups within a given schooling level. (3) The model uses an overlapping generations framework to produce heterogeneity among different cohorts as a result of rational investment behavior. In a period of transition, different skill price paths facing different entry cohorts produce important differences in the levels and rates of growth of earnings across cohorts. All three sources of heterogeneity are important in explaining rising wage inequality.

Our model considers human capital choices at both the extensive margin (schooling) and the intensive margin (on-the-job training). Schooling enables people to learn on the job and also directly produces market skills.

Our model extends the Roy [40] model of self-selection and earnings to allow for investment and embeds it in a dynamic general equilibrium model in which the prices of heterogenous skills are endogenously determined. It extends the widely used framework of Ben Porath [4] by permitting different technologies to govern the production of skill in schools and the production of skill on the job by recognizing that schooling affects both productivity on the job and the ability to learn on the job and by allowing for multiple skills. Our model extends the schooling models of Willis and Rosen [45] and Keane and Wolpin [27] by making post-schooling on-the-job training endogenous. It also extends those models and the analysis of Siow [42] by embedding both schooling and job training in a general equilibrium framework.

We relax the efficiency units assumption for the aggregation of labor services that is widely used in macroeconomics (see, e.g., [29]). This assumption is not consistent with rising wage inequality across skill groups except in the unlikely case where quantities of skill embodied in each group change over time in a fashion that exactly mimics movements in relative wages. Our model introduces human capital accumulation into the overlapping generations framework of Auerbach and Kotlikoff [1].[1]

In using our model to explain rising wage inequality in the American economy, we confront two major empirical problems, both of which are addressed in this paper. First, skills are unobserved. Second, we are not futurologists. We cannot forecast the progress of future technological change, nor do we know what forecasts of technological change agents act on. Yet agent expectations of future technical change play a critical role in any investment model, including our own.

To solve the first problem, we develop a new empirical procedure for estimating unobserved human capital stocks and use these stocks to produce the inputs required to estimate the parameters of aggregate technology and to test among competing specifications. Our methodology is of interest in its own right because it enables analysts to avoid the arbitrary measurement conventions about stocks of skills that are used in the empirical literature on estimating aggregate technology and in growth accounting.

Our estimated technology displays no skill bias between capital and skilled labor and produces an estimated elasticity of substitution between skilled and unskilled labor that is remarkably similar to previous estimates

---

[1]Davies and Whalley [12] were the first to introduce human capital into an overlapping generations model; however, they assume myopic expectations, assume a one-skill model, do not distinguish between schooling and on-the-job training, and ignore heterogeneity in endowments and human capital production technology within cohorts.

reported in the literature even though our measure of skill is fundamentally different from that used in previous work. We explain rising wage inequality without assigning any special role to capital.

To solve the second problem—that we do not know the expectations about future technology on which agents act—we proceed in the following way. Given our estimated ability-specific and skill-specific human capital investment equations and our estimated aggregate technology, we determine which assumptions about agent expectations of future technology produce a pattern of wage inequality consistent with the main features of the evidence.

We present new evidence on the sources of rising wage inequality in the U.S. economy. Unlike the models of Greenwood and Yorukoglu [14], Krussell et al. [28], Violante [44], and Caselli [10], our model explains why the recent (1979−1987) rise in wage inequality has been largest for new entrants among skilled workers and why age−earnings profiles have become steeper for less skilled workers and flatter for more skilled workers.[2]

An essential idea in our paper is that wages are not the same as prices —contrary to empirical conventions that equate the two. The link between skill prices and wages is broken by on-the-job investment, which is greatest for the young.[3] For them, the percentage gap between potential wages and measured wages is the greatest, because the greatest fraction of time is devoted to investment at early ages.

In response to a permanent upward shift in the demand for high-skilled labor, its price rises. This induces a supply response. More people go to college to obtain skills. At the beginning of the transition period, high-skill persons already working in the market respond by investing *more* on the job to take advantage of the unusually favorable future market for their services (given the parameters we estimate for the U.S. economy). Low-skill workers initially invest *less* on the job as their skill prices decline. They work as much as they can before the bottom drops out of their market. This initially depresses the college−high school wage differential. As the transition evolves, however, the roles of investment are reversed for the different skill groups. In the second phase, corresponding to a deceleration in the growth of the price of high-skill labor and a deceleration in the decline in the price of low-skill labor, the more skilled invest less and the less skilled invest more. Forward-looking investment behavior, therefore, causes wage differentials to overstate skill price differentials during this phase and contributes to greater measured short-run wage inequality. This could occur even if both groups of workers invested more in phase two, but

---

[2] See the evidence on these points in [26, Table I].

[3] This distinction is central to Mincer's [34] model of earnings, but has been neglected in recent studies of the determinants of earnings.

low-skill workers invest proportionately more than high-skill workers compared to their investment before the onset of technological change.

For cohorts born immediately after the shift in demand for skilled labor, college enrollment is higher than for all other cohorts. This creates a bulge in the supply of human capital that is like a price-induced baby boom. This creates a counterforce to rising skill prices and induces cycles in skill prices after technology stops changing.

We can explain rising wage inequality within educational groups, because persons of different ability levels respond differently to the same prices and, therefore, invest differently. With the responsiveness of the supply of skills to prices that is characteristic of recent entrant cohorts in the modern U.S. economy, we demonstrate that a demand shock biased toward skilled labor produces *lower* aggregate wage inequality in the long run as measured by the standard deviation of log wages. In the short run, the welfare of certain cohorts of low ability workers is substantially reduced, suggesting that their cases require special attention.

The disconnection between skill prices and wages due to skill investment decisions is a recurring theme of this paper. Because of this disconnection, the onset of skill-biased technical change can be masked. The wages of skilled workers may initially decline even though the price of skilled labor has increased. Uncritical use of wage data to signal demand or supply shifts is a dangerous but widespread practice.

We use our estimated model to determine whether immigration of unskilled workers into the U.S. economy plays a quantitatively important role in explaining rising wage inequality. While such migration can qualitatively account for rising wage inequality, the required magnitude of the immigration is far too large to make migration an important quantitative factor in our model.

We extend our model to explain the past 35 years of wage differentials in the United States. Augmenting the model to account for the enlarged cohorts of the Baby Boom, we find that the same estimated human capital production functions, aggregate production function, and estimated aggregate skill bias for technology used to explain rising wage inequality in the past 15 years can explain the rise, fall, and subsequent rise in the college−high school wage differential over the past 35 years and produces a rise in the variance of wages starting in the mid to late 1960s. Both of these patterns are generally consistent with wage patterns found in the U.S. data. This gives us confidence that we have isolated the structural parameters of human capital supply and aggregate technology for the U.S. economy.

The plan of this paper is as follows. Section I presents our general equilibrium framework. Section II presents the econometric framework used to identify parameters of the model presented in Section I and

presents estimates of our model. Section III uses the estimates to account for rising wage inequality. Section IV concludes the paper and discusses our proposed extensions.

I. PARTIAL EQUILIBRIUM AND GENERAL EQUILIBRIUM APPROACHES TO STUDYING WAGE INEQUALITY

A commonly used approach to assessing the sources of rising wage inequality starts with an equation postulated as a time-differenced demand relationship which connects the changes in the relative wages of skilled (W_{St}) and unskilled (W_{Ut}) workers at time t to the respective quantities of the two factors, Q_{St} and Q_{Ut}, respectively:

$$\Delta \ln\left[\frac{W_{St}}{W_{Ut}}\right] = \alpha - \frac{1}{\sigma}\Delta \ln\left[\frac{Q_{St}}{Q_{Ut}}\right]. \tag{I.1}$$

In this equation, α is the trend rate of relative wage growth arising from skill-biased technical change and σ is the elasticity of substitution between the two types of labor. Katz and Murphy [26] estimate this equation for the U.S. economy using the measures of skilled and unskilled labor defined in their paper for the period 1963–1987. They report $\sigma = 1.41$ with a standard error of 0.150, although they also suggest that a range of estimates with σ as low as $1/2$ are also consistent with the data.[4] They estimate α to be 0.033 (standard error 0.007).

Using their definition of skill groups, the 1979–1987 change in relative wages is roughly 0.14. A policy that reverses this trend by increasing the relative supply of skilled workers requires a once and for all increase of approximately 20% in the number of high-skill persons in the workforce. Using the Katz–Murphy definitions, college equivalents are 40% of the workforce and high school equivalents are 60%. For a 1990 workforce of 120 million, it is necessary to transform about 5.4 million people to college equivalents to reverse the decade-long erosion of real wages. Even using their lower range estimate of $\sigma = 0.5$, two million persons need to be shifted from the unskilled to the skilled category to offset the decade-long trend against unskilled labor.

To maintain this gain against the secular bias operating against unskilled labor, about one million additional skilled persons need to be added to the workforce each year (400 thousand for the lower bound case). As a

---

[4]Johnson [24] reported an estimate of $\sigma = 1.50$ for the elasticity of substitution between college and high school labor.

benchmark, the current annual supply of Katz–Murphy high-skill equivalents to the U.S. economy in the early 1990s is 1.8 million.[5] Maintaining the skill gaps requires that the percentage of persons acquiring post-secondary skills in each year rise by 55% (22% in the lower bound case).

Most policy evaluations that use micro estimates of the supply response of skills to tuition or other subsidies assume that skill prices remain constant at their pre-subsidy levels. They ignore the feedback of induced price changes created by the increase in the supply of skill on the supply decisions of agents. Only when these feedback responses are incorporated into supply decisions can Eq. (I.1) be used as a valid basis for policy evaluation. More convincing policy evaluations allow skill prices to adjust and agents to anticipate the adjustment and respond appropriately. Such evaluations recognize that the response to a policy evaluated in a microeconomic setting which holds prices constant may be a poor guide to the actual response when prices adjust in response to changes in quantities induced by the policy. Thus, for any policy characterized by parameter ψ, $\partial[Q_{S_t}/Q_{U_t}]/\partial\psi$ is not the same when skill prices are held fixed as it is when skill prices are allowed to vary in response to the policy-induced change.

A more convincing evaluation of any policy designed to promote skill formation to alleviate wage inequality requires a model of the sources of the rising wage inequality that is concordant with main features of the U.S. labor market. It is potentially dangerous to "solve" problems whose origins are not well understood. These concerns motivate us to develop a general-equilibrium model of labor earnings that is consistent with evidence from the U.S. labor market.

A. A Dynamic General Equilibrium Model of Earnings, Schooling, and On-the-Job Training

We generalize the microeconomic model of earnings, schooling, and on-the-job training developed by Ben Porath [4]. In his model, income maximizing agents combine time, goods, and the current stock of human capital to produce new human capital.

We extend Ben Porath's model in several ways. (1) In contrast to his model, we distinguish between schooling capital and job training capital at a given schooling level. In our model, schooling human capital is an input to the production of human capital acquired on the job and is also directly productive in the market. However, the tight link between schooling and on-the-job training investments which is characteristic of Ben Porath's model is broken. (2) Skills produced at different schooling levels command

---

[5]These calculations are presented in [19].

different prices, and wage inequality among persons is generated by differences in skill levels, differences in investment, and differences in the prices of different skills. In Ben Porath's model, wage inequality can only be generated by differences in skill levels and investment behavior, because all skill commands the same price. In our model, different levels of schooling enable individuals to invest in different skills through on-the-job training in the post-schooling period. In the aggregate, the skills associated with different schooling groups are not perfect substitutes.[6] Within schooling groups, however, persons with different amounts of skill are perfect substitutes.[7] Unlike the Ben Porath framework, our model of heterogeneous skills captures comparative advantage which is an important feature of modern labor markets.[8] (3) Persons choose among schooling levels with associated post-school investment functions. (4) Among persons of the same schooling level, there is heterogeneity both in initial stocks of human capital and in the ability to produce job-specific human capital. (5) We embed our model of individual human capital production into a general equilibrium setting so that the relationship between the capital market and the markets for human capital of different skill levels is explicitly developed. We extend the open-economy general equilibrium sectoral-choice model of Heckman and Sedlacek [22] to allow for investment in sector-specific human capital.

B. *The Microeconomic Model*

We first derive the optimal consumption, on-the-job investment, and schooling choices for a given individual of type θ who takes skill prices as given. We then aggregate the model to a general equilibrium setting. Throughout this paper, we simplify the tax code and assume that income taxes are proportional.[9] Individuals live for $\bar{a}$ years and retire after $a_R \leq \bar{a}$ years. Retirement is mandatory. In the first portion of the life cycle, a prospective student decides whether or not to remain in school. Once he has left school, he cannot return. He chooses the schooling option that gives him the highest level of lifetime utility.

Define K_{at} as the stock of physical capital held at time t by a person age a; H_{at}^S is the stock of human capital at time t of type S at age a. The optimal life-cycle problem can be solved in two stages. First, condition on

---

[6]This specification is consistent with evidence that the large increase in the supply of educated labor consequent from the Baby Boom depressed the returns to education (see [13, 2, 26]).

[7]This specification accords with the empirical evidence summarized in [15, p. 123] that persons of different ages but with the same education levels are highly substitutable for each other.

[8]See the empirical evidence summarized in [41].

[9]In other work this assumption is relaxed; see [21].

schooling and solve for the optimal path of consumption (C_{at}) and post-school investment (I_{at}^S) for each type of schooling level. Individuals then select among schooling levels to maximize lifetime welfare. Given S, an individual age a at time t has the value function

$$V_{at}(H_{at}^S, K_{at}, S) = \max_{C_{at}, I_{at}^S} U(C_{at}) + \delta V_{a+1,t+1}(H_{a+1,t+1}^S, K_{a+1,t+1}, S),$$

(I.2)

where U is strictly concave and increasing and δ is a time preference discount factor. This function is maximized subject to the budget constraint

$$K_{a+1,t+1} \le K_{a,t}(1 + (1 - \tau)r_t) + (1 - \tau)R_t^S H_{at}^S(1 - I_{at}^S) - C_{at}, \quad (I.3)$$

where τ is the proportional tax rate on capital and labor earnings, R_t^S is the rental rate on human capital of type S, and r_t is the net return on physical capital at time t. In this paper, we abstract from labor supply. Estimates of intertemporal substitution in labor supply estimated on annual data are small, so ignoring labor supply decisions will not greatly affect our analysis.[10]

In the empirical analysis in this paper, we use the conventional power utility specification of preferences

$$U(C_{at}) = \frac{C_{at}^\gamma - 1}{\gamma}.$$

On-the-job human capital for a person of schooling level S accumulates through the human capital production function

$$H_{a+1,t+1}^S = A^S(\theta)(I_{a,t}^S)^{\alpha_S}(H_{a,t}^S)^{\beta_S} + (1 - \sigma^S)H_{a,t}^S, \quad (I.4)$$

where the conditions $0 < \alpha_S < 1$ and $0 \le \beta_S \le 1$ guarantee that the problem is concave in the control variable, and σ^S is the rate of depreciation of job-S-specific human capital. This functional form is widely used in both the empirical literature and the literature on human capital accumulation.[11]

For simplicity, we ignore the input of goods into the production of human capital on the job. For an analysis of post-school investment, this is

[10] See [7] or the survey in [18].

[11] Uzawa [43] assumed $\beta_S = 1$. Ben Porath [4], [31] and Ortigueira and Santos [36] assumed that $\alpha_S = \beta_S$. Rosen [39] assumed $\alpha_S = 1/2$ and $\beta_S = 1$.

not restrictive as we can always introduce goods and solve them out as a function of I_{at}^S, thereby reinterpreting I_{at}^S as a goods–time investment composite. We explicitly allow for tuition costs of college which we denote by D_{at}^S. The same good that is used to produce capital and final output is used to produce schooling human capital. After completion of schooling, time is allocated to two activities, both of which must be nonnegative: on-the-job investment, I_{at}^S, and work, $(1 - I_{at}^S)$. The agent solves a life-cycle optimization problem given initial stocks of human and physical capital, $H^S(\theta)$ and K_0, as well as his ability to produce human capital on the job, $A^S(\theta)$.

$H^S(\theta)$ and $A^S(\theta)$ represent ability to ''earn'' and ability to ''learn,'' respectively, measured *after* completion of school. They embody the contribution of schooling to subsequent learning and earning in the schooling-level S-specific skills as well as any initial endowments. Notably absent from our model are the short-run credit constraints that are often featured in the literature on schooling and human capital accumulation. Our model is consistent with the evidence presented in [8, 9] that long-run family factors correlated with income (the θ operating through $A^S(\theta)$ and $H^S(\theta)$) affect schooling, but that short-term credit constraints are not empirically important. Such long-run factors account for the empirically well-known correlation between schooling attainment and family income. The mechanism generating the family income–schooling relationship operates through family-acquired human capital and not credit rationing. The α and β are also permitted to be S-specific, which emphasizes that schooling affects the process of learning on the job in a variety of different ways.

Assuming interior solutions conditional on the choice of schooling, we obtain the following first order conditions:

$$U_{C_{a,t}} = \delta \frac{\partial V_{a+1,t+1}}{\partial K_{a+1,t+1}} (1 - \tau), \tag{I.5}$$

$$\frac{\partial V_{a+1,t+1}}{\partial K_{a+1,t+1}} = \delta \frac{\partial V_{a+1,t+1}}{\partial H_{a+1,t+1}} \left[\frac{A \alpha_S \left(I_{a,t}^S \right)^{\alpha_S - 1} \left(H_{a,t}^S \right)^{\beta_S}}{R_t^S H_{a,t}^S (1 - \tau)} \right]$$

(*marginal return to investment time equals marginal cost*)   (I.6)

$$\frac{\partial V_{a,t}}{\partial K_{a,t}} = \delta \frac{\partial V_{a+1,t+1}}{\partial K_{a+1,t+1}} (1 + r_t(1 - \tau))$$

(*intertemporal arbitrage in returns on physical capital*)   (I.7)

$$\frac{\partial V_{a,t}}{\partial H_{a,t}^S} = \delta \frac{\partial V_{a+1,t+1}}{\partial K_{a+1,t+1}} R_t^S (1 - I_{a,t}^S)(1 - \tau)$$

$$+ \delta \frac{\partial V_{a+1,t+1}}{\partial H_{a+1,t+1}^S} \left(A \beta_S (I_{a,t}^S)^{\alpha_S} (H_{a,t}^S)^{\beta_S - 1} + (1 - \delta^S) \right)$$

(*marginal value of human capital is the return to current and future earnings*).

$$(I.8)$$

At the end of working life, the final term, which is the contribution of human capital to earnings, has zero marginal value. We assume mandatory retirement at age a_R, leaving ages $\bar{a} - a_R$ as the retirement period during which there are no labor earnings.

At the beginning of life, agents choose the value of S that maximizes lifetime utility,

$$\hat{S} = \underset{S}{\text{Arg max}} \left[V^S(\theta) - D^S - \varepsilon^S \right], \tag{I.9}$$

where $V^S(\theta)$ is the present value of earnings for schooling at level S, D^S is the discounted direct cost of schooling, and ε^S represents nonpecuniary benefits expressed in present value terms. Discounting of V^S and D^S is back to the beginning of life to account for different ages of completing school. Tuition costs are permitted to change over time so that different cohorts face different environments for schooling costs. Given optimal investment in physical capital, schooling, investment in job-specific human capital, and consumption, we calculate the path of savings. For a given return on capital and rental rates on human capital, the solution to the S-specific optimization problem is unique given concavity of the production function of (I.4) in terms of I_{at}^S ($0 < \alpha_S < 1$), the restriction that human capital be self-productive, but not too strongly ($0 \le \beta_S \le 1$), the restriction that investment is in the unit interval ($0 \le I_{at}^S \le 1$), and concavity of U in terms of C ($\gamma < 1$).

The choice of S is unique almost surely if ε^S is a continuous random variable, as we assume in our empirical analysis. The dynamic problem is of split-endpoint form. We know the initial condition for human and physical capital and optimality implies that investment is zero at the end of life. In this paper, we numerically solve this problem using the method of "shooting" (see [30]). For any terminal value of H^S and K, we solve backward to the initial period and obtain the implied initial conditions. We iterate until the simulated initial condition equals the prespecified value.

C. Aggregating the Model

The prices of skills and capital are determined as derivatives of an aggregate production function. To compute rental prices for capital and the different types of human capital, it is necessary to construct aggregates of each of the skills. Given the solution to the individual's problem for each value of θ and each path of prices, we use the distribution of θ, $G(\theta)$, to construct aggregates of human and physical capital. We embed our human capital model into an overlapping generations framework in which the population at any given time is composed of $\bar{a}$ overlapping generations, each with an identical ex-ante distribution of heterogeneity, $G(\theta)$.

Human capital of type S is a perfect substitute for any other human capital of the same schooling type, whatever the age or experience level of the agent, but it is not perfectly substitutable with human capital from other schooling levels. In our model, cohorts differ from each other only because they face different price paths and policy environments within their lifetimes. We assume perfect foresight (as used in [1]) and not myopic expectations. Let c index cohorts, and denote the date at which cohort c is born by t_c. Their first period of life is $t_c + 1$. Let P_{t_c} be the vector of paths of rental prices of physical and human capital confronting cohort c over its lifetime from time $t_c + 1$ to $t_c + \bar{a}$. The rental rate on physical capital at time t is r_t. The rental rate on human capital is R_t^S. The choices made by individuals depend on the prices they face, P_{t_c}, their type, θ, and hence their endowment and their nonpecuniary costs of schooling, ε^S. Let $H_{at}^S(\theta, P_{t_c})$ and $K_{at}^S(\theta, P_{t_c})$ be the amount of human and physical capital possessed, respectively, and let $I_{at}^S(\theta, P_{t_c})$ be the time devoted to investment by an individual with schooling level S, at age a, of type θ, in cohort c.

By definition, the age at time t of a person born at time t_c is $a = t - t_c$. Let $N^S(\theta, t_c)$ be the number of persons of type θ, in cohort c, of schooling level S. In this notation, the aggregate stock of employed human capital of type S at time t is cumulated over the nonretired cohorts in the economy at time t,

$$\bar{H}_t^S = \sum_{t_c = t - a_R}^{t-1} \int H_{t-t_c,t}^S(\theta, P_{t_c})\left(1 - I_{t-t_c,t}^S(\theta, P_{t_c})\right) N^S(\theta, t_c)\, dG(\theta),$$

where $a = t - t_c$, $S = 1, \ldots, \bar{S}$, where $\bar{S}$ is the maximum number of years of schooling. The aggregate *potential* stock of human capital of type S is

obtained by setting $I_a^S(\theta, P_{t_c}) = 0$ in the preceding expression:

$$\overline{H}_t^S(\text{potential}) = \sum_{t_c=t-a_R}^{t-1} \int H_{t-t_c,t}^S(\theta, P_{t_c}) N^S(\theta, t_c)\, dG(\theta).$$

The aggregate capital stock is the capital held by persons of all ages:

$$\overline{K}_t = \sum_{t_c=t-\overline{a}}^{t-1} \sum_{s=1}^{\overline{S}} \int K_{t-t_c,t}^S(\theta, P_{t_c}) N^S(\theta, P_{t_c})\, dG(\theta).$$

D. Equilibrium Conditions under Perfect Foresight

To close the model, it is necessary to specify the aggregate production function $F(\overline{H}_t^1, \ldots, \overline{H}_t^S, \overline{K}_t)$, which is assumed to exhibit constant returns to scale. The equilibrium conditions require that marginal products equal pre-tax prices $R_t^S = F_{\overline{H}_t^S}(\overline{H}_t^1, \ldots, \overline{H}_t^{\overline{S}}, \overline{K}_t, t)$, $S = 1, \ldots, \overline{S}$, and $r_t = F_{\overline{K}_t}(\overline{H}_t^1, \ldots, \overline{H}_t^{\overline{S}}, \overline{K}_t, t)$. In the two-skill economy estimated below, we specialize the production function to

$$F\left(\overline{H}_t^1, \overline{H}_t^2, \overline{K}_t\right)$$
$$= a_3 \left(a_2 \left(a_1 \left(\overline{H}_t^1\right)^{\rho_1} + (1-a_1)\left(\overline{H}_t^2\right)^{\rho_1} \right)^{\rho_2/\rho_1} + (1-a_2) \overline{K}_t^{\rho_2} \right)^{1/\rho_2}. \tag{I.10}$$

When $\rho_1 = \rho_2 = 0$, the technology is Cobb−Douglas.[12] When $\rho_2 = 0$, we obtain a model consistent with the constancy of capital's share irrespective of the value of ρ_1.

Activities of the government, apart from its role in subsidizing human capital, are not central to our analysis. The government collects taxes at a fixed level and does not redistribute them.

E. Linking the Earnings Function to Prices and Market Aggregates

The earnings for a person of age a of cohort c of type θ with human capital $H_{a,t}^S(\theta, P_{t_c})$ at time t are

$$W\left(a, t, H_{a,t}^S(\theta, P_{t_c})\right) = R_t^S H_{a,t}^S(\theta, P_{t_c})\left(1 - I_{a,t}^S(\theta, P_{t_c})\right). \tag{I.11}$$

---

[12]Auerbach and Kotlikoff [1] assumed efficiency units so different labor skills are perfect substitutes ($\rho_1 = 1$). In addition, they assumed a Cobb−Douglas aggregate technology relating human capital and physical capital ($\rho_2 = 0$).

They are determined by aggregate rental rates (R_t^S), individual endowments $H_{at}^S(\theta, P_{t_c})$, and individual investment decisions $(I_a^S(\theta, P_{t_c}))$ The last two components depend on agent expectations of future prices. Different cohorts facing different price paths will invest differently and have different human capital stocks. An essential idea in this paper, which is absent from currently used specifications of earnings equations in labor economics, is that *utilized* skills and not potential skills determine earnings. The utilization rate is an object of choice linked to personal investment decisions and is affected both by individual endowments and aggregate skill prices. As the quantity of aggregate skill is changed, so are aggregate skill prices. This affects investment decisions, measured wages, and savings decisions.

II. DETERMINING THE PARAMETERS OF THE MODEL

This section discusses how we choose the parameters of our model. In estimating skill-specific human capital production functions, we account for individual heterogeneity in technology and endowments. We present a new test of the consistency of the estimates of model parameters used to generate the general equilibrium simulations. We require that the econometric procedure used to produce the micro-based parameters employed in our model (including the implicit assumptions made about the economic environment in implementing any particular econometric procedure) recover the parameters estimated from synthetic micro data sets generated by the model used to simulate the economy. We further require that our assumptions about agent expectations produce the behavior observed in our sample period.

The ideal data set for our purposes would combine micro data on firms, data on the earnings of workers, their life-cycle consumption, and their wealth holdings, and macro data on prices and aggregates. With such data, we could estimate all the parameters of our model and the distribution of wages, wealth, and earnings. Using the micro data joined with aggregate prices, we could estimate the parameters of the micro model. Using the estimated micro functions, we could construct aggregates of human capital that can be used in determining the output technology. The estimated aggregates should match measured empirical aggregates and, when inserted in aggregate technology, should also reproduce the market prices used in estimation.

Two obstacles prevent us from implementing this approach. (1) We lack information on individual consumption linked to labor earnings. (2) The data on market wages do *not* reveal skill prices, as is evident from the distinction between R_t^S and $W(a, t, H_{a,t}^S(\theta, P_{t_c}))$ in Eq. (I.11). Since prices

cannot be directly equated with wages, it is apparently not possible to estimate aggregate stocks of human capital to use in determining aggregate technology.

To circumvent the first limitation, we follow practices widely used in the literature on empirical general equilibrium models, and we choose discount and intertemporal substitution parameters in consumption to be consistent with those reported in the empirical literature and that enable us to reproduce key features of the macro data—like the capital–output ratio. We explore the sensitivity of our simulations to alternative choices of these parameters.

To circumvent the second problem, we develop a new method for using wages to infer prices and to estimate skill-specific human capital aggregates. Since calibration methods and sensitivity analysis are widely used in applied general equilibrium analysis, we turn to the more original empirical contribution of this paper.

A. Simple Methods for Estimating Skill Prices and Aggregate Production Technology with Heterogeneous Skills

We first present a method for identifying the aggregate technology and estimating skill–specific human capital stocks by combining micro and macro data. It exploits the insight that at older ages, changes in wages are due solely to changes in skill prices and to depreciation. Suppose that for two consecutive ages, a and $a + 1$, $I_{a,t}^S = I_{a+1,t+1}^S = 0$. More ages of zero investment only help to identify skill prices, so we present a worst-case analysis. At late ages in the life cycle, $I_{a,t}^S \cong 0$ is an implication of optimality. This condition enables us to identify rental rates up to scale. Note that from the definitions (dropping the θ and c subscripts for simplicity) and from the identifying assumption at ages a and $a + 1$, for older cohorts it follows that

$$W\left(a + 1, t + 1, H_{a+1,t+1}^S\right) \equiv R_{t+1}^S H_{a+1,t+1}^S = R_{t+1}^S H_{a,t}^S (1 - \sigma^S),$$

where σ^S is the rate of depreciation for skill S. It is assumed that deflated real wages are used. Then it follows that

$$\left(\frac{W\left(a + 1, t + 1, H_{a+1,t+1}^S\right)}{W\left(a, t, H_{a,t}^S\right)} \right) = \left(\frac{R_{t+1}^S(1 - \sigma^S)}{R_t^S} \right).$$

Normalize $R_0 = 1$. In the absence of measurement error in wage ratios, we can identify $R_0^S, \ldots, R_T^S$ from a time series of cross sections of individuals ages a and $a + 1$ up to scale $(1 - \sigma^S)^t$, where t is the time period; i.e., we can identify $R_0^S, R_1^S(1 - \sigma^S), \ldots, R_T^S(1 - \sigma^S)^T$. If the ratios have mean

zero measurement error, we can estimate the ratios of skill prices without bias. With these relative skill prices in hand, we can recover utilized human capital stocks up to scale.

Denote WB_t^S as the total wage bill in the economy at time t for schooling level S. This is available from the aggregate information in a time series of cross sections on wages. Then it is possible to estimate the aggregate utilized human capital stock of type S up to scale from the equation

$$\frac{\mathrm{WB}_t^S}{(1 - \sigma^S)^t R_t^S} = \frac{\sum_{a=1}^{a_R} H_{a,t}^S (1 - I_{a,t}^S)}{(1 - \sigma^S)^t} = \frac{\overline{H}_t^S}{(1 - \sigma^S)^t}$$

so that we can generate human capital stocks at time t up to scale $(1 - \sigma^S)^t$. We now show how to identify σ^S, and recover the aggregate technology. For later use, define the measured variables by the tilde ($\tilde{\ }$) and write $\tilde{R}_t^S = R_t^S (1 - \sigma^S)^t$ and $\tilde{H}_t^S = \overline{H}_t^S / (1 - \sigma^S)^t$.

B. Identifying σ^S and the Aggregate Technology

From aggregate production technology (I.10) and the assumption of market clearing in competitive markets, we obtain the following first order conditions which generate skill prices. To simplify the derivation, we first define

$$Q_t = \left[a_1 \left(\overline{H}_t^1 \right)^{\rho_1} + (1 - a_1) \left(\overline{H}_t^2 \right)^{\rho_1} \right]^{1/\rho_1}$$

so the aggregate technology can be written in terms of the composite

$$F\left(\overline{H}_t^1, \overline{H}_t^2, \overline{K}_t \right) = \left[(1 - a_2) Q_t^{\rho_2} + a_2 \overline{K}^{\rho_2} \right]^{1/\rho_2}.$$

Let $C_t = [(1 - a_2) Q_t^{\rho_2} + a_2 \overline{K}_t^{\rho_2}]^{(1 - \rho_2)/\rho_2}$. Then in this notation

$$r_t = C_t a_2 \overline{K}_t^{\rho_2 - 1},$$

$$R_t^1 = C_t (1 - a_2) a_1 Q_t^{\rho_2 - 1} \left(\overline{H}_t^1 \right)^{\rho_1 - 1},$$

$$R_t^2 = C_t (1 - a_2)(1 - a_1) Q_t^{\rho_2 - 1} \left(\overline{H}_t^2 \right)^{\rho_1 - 1}.$$

The log ratio of the last two optimality conditions is

$$\log\left(\frac{R_t^2}{R_t^1}\right) = \log\left(\frac{1-a_1}{a_1}\right) + (\rho_1 - 1)\log\left(\frac{\overline{H}_t^2}{\overline{H}_t^1}\right).$$

In terms of the measured variables denoted by the tilde, this equation can be written as

$$\log\left(\frac{\tilde{R}_t^2}{\tilde{R}_t^1}\right) = \log\left(\frac{1-a_{1t}}{a_{1t}}\right) + \rho_1\log\left(\frac{1-\sigma^2}{1-\sigma^1}\right)t + (\rho_1 - 1)\log\left(\tilde{H}_t^2/\tilde{H}_t^1\right),$$

(II.1)

where we permit the a_1 to depend on time. We allow linear trends in $\log((1-a_1)/a_1)$, so $\log[(1-a_{1t})/a_{1t}] = \log[(1-a_{10})/a_{10}] + \varphi t$, where $t = 0$ is the baseline period. We may write this expression in terms of the relative wage bill for skill S at time t, denoted by WB_t^S:

$$\log\left(\text{WB}_t^2/\text{WB}_t^1\right) = \log[(1-a_{10})/a_{10}]$$
$$+ \left[\varphi + \log[(1-\sigma^2)/(1-\sigma^1)]\right]t + \rho_1\log\left(\tilde{H}_t^2/\tilde{H}_t^1\right).$$

Ordinary least squares applied to

$$\log\left(\frac{\tilde{R}_t^2}{\tilde{R}_t^1}\right) = \alpha + \beta\log\left(\frac{\tilde{H}_t^2}{\tilde{H}_t^1}\right) + \eta t + \varepsilon_t \qquad (\text{II.2})$$

consistently estimates $\alpha = \log[(1-a_{10})/a_{10}]$, $\beta = \rho_1 - 1$, $\eta = \rho_1\log[(1-\sigma^2)/(1-\sigma^1)] + \varphi$ if there are measured shifters that are exogenous to ε_t. In this case, the ratio of depreciation rates $(1-\sigma^2)/(1-\sigma^1)$ is identified if the technology is stable ($\varphi = 0$).

To recover the other parameters, observe that from CES algebra the price of the bundle Q is

$$R_t^Q = \left((R_t^1)^{\rho_1/(\rho_1-1)}(a_1)^{1/(1-\rho_1)} + (R_t^2)^{\rho_1/(\rho_1-1)}(1-a_1)^{1/(1-\rho_1)}\right)^{(\rho_1-1)/\rho_1}.$$

Thus we may write the first order condition for Q and K as

$$\log\left(\frac{R_t^Q}{r_t}\right) = \log\left(\frac{a_2}{1-a_2}\right) + (\rho_2 - 1)\log\left(\frac{Q_t}{\overline{K}_t}\right). \qquad (\text{II.3})$$

In terms of observables, or variables that can be derived from observables,

$$\tilde{R}_t^Q = \left[\left(\tilde{R}_t^1 \right)^{\rho_1/(1-\rho_1)} (a_1)^{1/(1-\rho_1)} (1 - \sigma^1)^{-[\rho_1/(\rho_1-1)]t} \right.$$

$$\left. + \left(\tilde{R}_t^2 \right)^{\rho_1/(1-\rho_1)} (1 - a_1)^{1/(1-\rho_1)} (1 - \sigma^2)^{-[\rho_1/(\rho_1-1)]t} \right]^{(\rho_1-1)/\rho_1}.$$

In the special case where $\sigma^1 = \sigma^2 = \sigma$ and a_1 is constant over time, we may write this expression as

$$\tilde{R}_t^Q = (1 - \sigma)^{-t} \left(\left(\tilde{R}_t^1 \right)^{\rho_1/(1-\rho_1)} (a_1)^{1/(1-\rho_1)} \right.$$

$$\left. + \left(\tilde{R}_t^2 \right)^{\rho_1/(1-\rho_1)} (1 - a_1)^{1/(1-\rho_1)} \right)^{\rho_1/(1-\rho_1)}.$$

Similarly,

$$\tilde{Q}_t = \left[a_1 \left(\tilde{H}_t^1 \right)^{\rho_1} (1 - \sigma^1)^{t\rho_1} + (1 - a_1) \left(\tilde{H}_t^2 \right)^{\rho_1} (1 - \sigma^2)^{t\rho_1} \right]^{1/\rho_1}.$$

Again, when $\sigma^1 = \sigma^2 = \sigma$,

$$\tilde{Q}_t = (1 - \sigma)^t \left[a_1 \left(\tilde{H}_t^1 \right)^{\rho_1} + (1 - a_1) \left(\tilde{H}_t^2 \right)^{\rho_1} \right]^{1/\rho_1}.$$

For this special case, we may write the estimating equation for (II.3) as

$$\log\left(\frac{\tilde{R}_t^Q}{r_t} \right) = \log\left(\frac{a_2}{1 - a_2} \right) + \left[\rho_2 \ln(1 - \sigma) \right]t + (\rho_2 - 1)\log\left(\frac{\tilde{Q}_t}{\overline{K}_t} \right) + \nu_t. \tag{II.4}$$

In the general case, with $\sigma^1 \neq \sigma^2$, it is necessary to use nonlinear least squares to construct the aggregates and prices conditional on σ^1 and σ^2 and to iterate until a best fitting model is found. Instruments are required if there are demand shocks.[13] A test of $\rho_2 = \rho_1 = 0$ is a test of the Cobb–Douglas specification for aggregate technology using the constructed skill prices and aggregates obtained from the first stage estimation procedure.

_____

[13] Observe that in forming R_t^Q and using it in subsequent estimation, one should correct for parameter estimation error in forming $\tilde{Q}_t$ in order to produce correct standard errors.

C. Estimating Human Capital Production Functions

We estimate human capital production parameters using NLSY data on white male earnings for the period 1979–1993. We follow the literature [16, 17] and assume that interest rates and the after-tax rental rates on human capital are fixed at constant but empirically concordant values. This ignores the price variation induced by technological change. A remarkable finding of our research, reported in Appendix B, is that this misspecification of the economic environment has only slight consequences for the estimation of the curvature parameters of the human capital technology, at least within the range of variation of skill prices generated by our model for the U.S. economy. Misspecification in share parameters is compensated for by calibration.

We take the real after-tax interest rate r as given and fix it at 0.05 which is in the range of estimates reported by Poterba [38] for our sample period. We treat R_t^S as a constant (normalized to 1) for all skill services following a tradition in the literature. We set $\sigma^S = 0$, an estimate consistent with what is reported in the literature (see [7]). It is consistent with the lack of any peak in life-cycle wage–age profiles reported in the literature (see [33]). We use a tax rate of 15% which is consistent with the effective rate over our sample period reported by Pechman [37]. For each ability–schooling (θ, S) type, the relevant parameters are $(\alpha^S, \beta^S, A^S(\theta), H_0^S(\theta))$. We assume that, conditional on measured ability, there is no dependence in unobservables across the schooling and wage equations.

Solving the human capital model backward is easier computationally than simultaneously solving it forward and backward. Rather than parameterizing the model in terms of initial human capital, we parameterize it in terms of terminal human capital, denoted $H_{a_R}^S(\theta)$. Since there is a one-to-one relationship between initial human capital and terminal human capital, this parameterization is innocuous.

For any particular set of parameters $(\alpha^S, \beta^S, A^S(\theta), H_{a_R}^S(\theta))$, we can simulate the model and form log wage profiles as functions of these parameters. For each ability type (θ), we estimate the model by nonlinear least squares minimizing over individuals (denoted i),

$$\sum_i \sum_a \left(W_{i,a}^* - W_a\left(\alpha^S, \beta^S, A_\theta^S(\theta), H_{a_R}^S(\theta)\right) \right)^2,$$

and constraining $0 < \alpha_S < 1$, $0 \leq \beta_S \leq 1$, and $A^S(j) > 0$ for two schooling groups. $S = 2$ if a person has completed one year of college; $S = 1$ otherwise. Estimates from this model are presented in the first panel of Table I. Level $\theta = 1$ is the lowest quartile of AFQT ability while $\theta = 4$ is the highest quartile. The estimates of α^S and β^S are quite similar for the

TABLE I
Estimated Parameters for Human Capital Production Function
and Schooling Decision[a]
(standard errors in parentheses)

| | Human capital production $H_{a+1}^S = A^S(\theta)I_a^{\alpha_S}H_a^{\beta_S} + (1 - \sigma)H_a^S; \; S = 1, 2$ | |
	High school $(S = 1)$	College $(S = 2)$
α	0.945 (0.017)	0.939 (0.026)
β	0.832 (0.253)	0.871 (0.343)
$A(1)$	0.081 (0.045)	0.081 (0.072)
$H_{a_R}(1)$	9.530 (0.309)	13.622 (0.977)
$A(2)$	0.085 (0.053)	0.082 (0.074)
$H_{a_R}(2)$	12.074 (0.403)	14.759 (0.931)
$A(3)$	0.087 (0.056)	0.082 (0.077)
$H_{a_R}(3)$	13.525 (0.477)	15.614 (0.909)
$A(4)$	0.086 (0.054)	0.084 (0.083)
$H_{a_R}(4)$	12.650 (0.534)	18.429 (1.095)

| | College choice equation $P(\delta^2 = 1) = \Lambda(-\lambda D^2 + \alpha(\theta))$ | |
	Probit parameters	Average derivatives
λ	0.166 (0.062)	-0.0655 (0.025)
$\alpha(1)$	-1.058 (0.097)	—
$\alpha(2)$	-0.423 (0.087)	0.249 (0.037)
$\alpha(3)$	0.282 (0.089)	0.490 (0.029)
$\alpha(4)$	1.272 (0.101)	0.715 (0.018)
Sample size		
Persons	869	1,069
Person Years	7,996	11,626

[a] D^2 is the discounted tuition cost of attending college. $\alpha(\theta)$ is the nonparametric estimate of $(1 - \tau)[V^2(\theta) - V^1(\theta)]$, the monetary value of the gross discounted returns to attending college. $\delta^2 = 1$ if attend college; $\delta^2 = 0$ otherwise. Λ is the unit normal cdf.

two schooling groups. The value of the productivity parameters $A^S(\theta)$ usually increase with AFQT, suggesting that more able people are more efficient in producing human capital. The terminal levels of human capital are higher for college-educated individuals than for persons attending high school.

In Appendix B, we present a sensitivity analysis of our estimates of the model to misspecification of heterogeneity and the economic environment. Except for the case where we estimate the model under one interest rate and simulate the model under another, the model is surprisingly robust to misspecifications of the economic environment. We find that when skill prices are varying in the economy but are assumed fixed in our econometric procedure, estimates of the curvature parameters of the human capital production function are barely affected, at least within the range of prices produced by the simulations of our model.

In the first panel of Table II, we present the initial levels of human capital for each schooling type by solving the model backward given the terminal condition estimates reported in Table I. For all four ability types, the job market entry level of human capital increases with college. Except for one case, the initial level of human capital increases with AFQT across schooling groups. The one anomaly in the table is that the initial level of human capital is larger for the high school group of ability type 3 than for

TABLE II

Derived Parameters for Human Capital Production Function and Schooling Decision[a]

(units are thousands of dollars)

	Human capital production	
	High school ($S = 1$)	College ($S = 2$)
$H^S(1)$	8.042 (0.094)	11.117 (0.424)
$H^S(2)$	10.063 (0.118)	12.271 (0.325)
$H^S(3)$	11.127 (0.155)	12.960 (0.272)
$H^S(4)$	10.361 (0.234)	15.095 (0.323)
Present value earnings 1	260.304 (3.939)	289.618 (12.539)
Present value earnings 2	325.966 (5.075)	319.302 (10.510)
Present value earnings 3	360.717 (6.352)	337.260 (9.510)
Present value earnings 4	335.977 (8.453)	393.138 (11.442)

College decision: attend college if $(1 - \tau)V^2(\theta) - D^2 + \varepsilon_i \geq (1 - \tau)V^1(\theta)$; $\varepsilon_\theta \sim N(\mu_\theta, \sigma_\varepsilon)$

σ_ε (Std. deviation of ε)	22.407 (8.425)
Nonpecuniary costs by ability level	
μ_1 (lowest ability quartile)	-53.019 (16.770)
μ_2 (second ability quartile)	-2.817 (12.760)
μ_3 (third ability quartile)	29.771 (11.540)
μ_4 (highest ability quartile)	-28.649 (16.966)

[a] $V^i(\theta)$ is the monetary value of going to schooling level i for a person of AFQT quartile θ. $i = 1$ for high school; $i = 2$ for college. ε_θ is the nonpecuniary benefit of attending college for a person of ability quartile θ. D^2 is the discounted tuition cost of attending college.
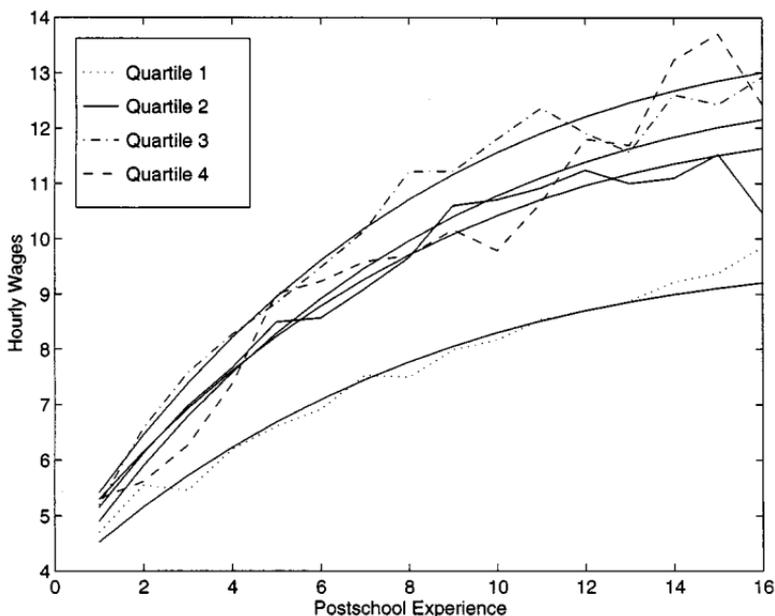
FIG. 1. Predicted vs actual hourly wages (in 1992 dollars) by AFQT quartile (high school category).

the high school group of ability type 4. (This is consistent, however, with the earnings profiles of young men in the NLSY displayed in Fig. 1.) The present value of earnings increases with ability for most groups.

The estimates of the human capital production function reported in Tables I and II are consistent with the Ben Porath model ($\alpha^S = \beta^S$) that was widely used in the early literature on estimating human capital technology. The point estimates are remarkably similar to those reported by Heckman [17] for his income maximizing models ($\alpha^S = \beta^S = 0.812$) for males and are consistent with the range of estimates reported by Brown [6] for his sample of young males.[14] The estimated models fit the earnings data rather well for different schooling and ability levels. See Figs. 1 and 2.

The implied investment profile declines much more steeply than the one assumed by Mincer [34]. (See Fig. 3.) Mincer assumed that $I^S(\tau) = k^S(1 - \tau/T)$, where τ is post-school work experience, T is the effective working life, and k^S ($0 \leq k^S < 1$) is a constant assumed to be the same at all levels of schooling. Our estimates reveal a much sharper decline with experience

---

[14]Our estimates for women are quite similar to those for males [21]. For them "β^S" and "α^S" are high and we cannot reject the Ben Porath model as a description of their human capital production function, or that women and men have the same human capital production function.
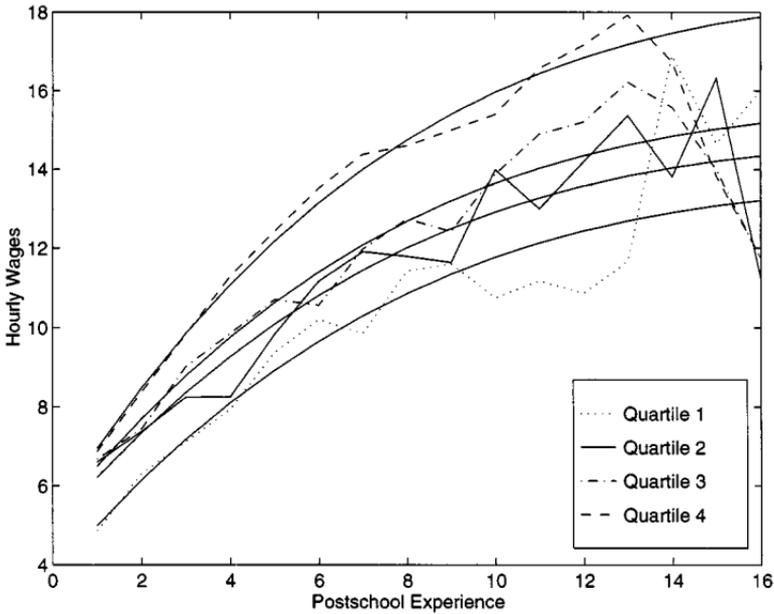
FIG. 2. Predicted vs actual hourly wages (in 1992 dollars) by AFQT quartile (college category).

and also indicate that k^S *increases* with schooling.[15] This produces divergence in log wage–experience profiles among schooling levels with the more skilled experiencing more rapid wage growth. The interaction is small, however, and in the data the departure from log parallelism in earnings profiles is slight [23].

Using our model and the assumption of no depreciation in skills, we can estimate the contribution of schooling and on-the-job training to lifetime human capital. Using an accounting framework that equates marginal and average rates of return, Mincer [34] estimates that half of all human capital formation is on the job. Using our optimizing framework, which distinguishes between marginal and average rates of return, we find that the contribution of OJT to the total human capital stock is much less—on the order of 23%—over all ability groups.

D. Estimating the Probability of Attending College

Let D_i be the college tuition faced by individual i. The difference in utility between going to college or not for individual i who is of ability type

---

[15] The solution for Mincer's k^S is obtained from the coefficients of estimated experience in a regression of simulated earnings on S and experience squared assuming $T = 47$ years for high school and $T = 43$ years for college students using the formulae in Mincer [35].
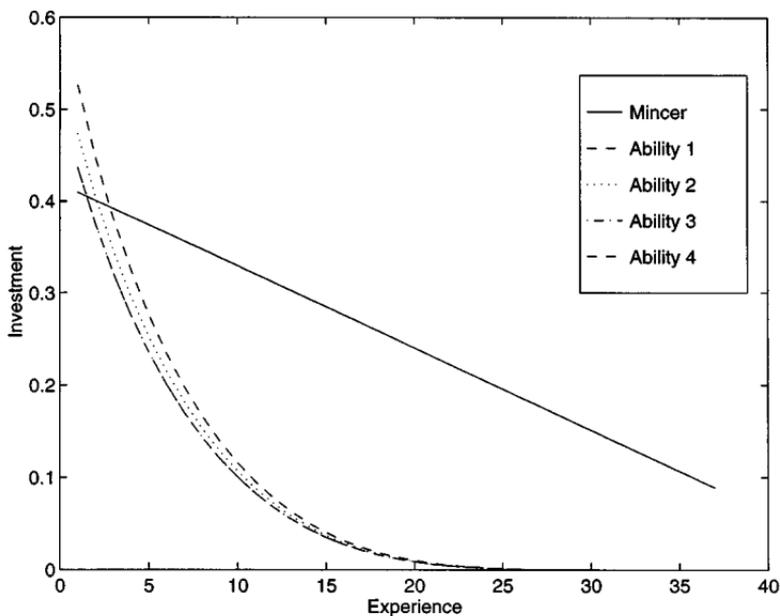
(a) high school
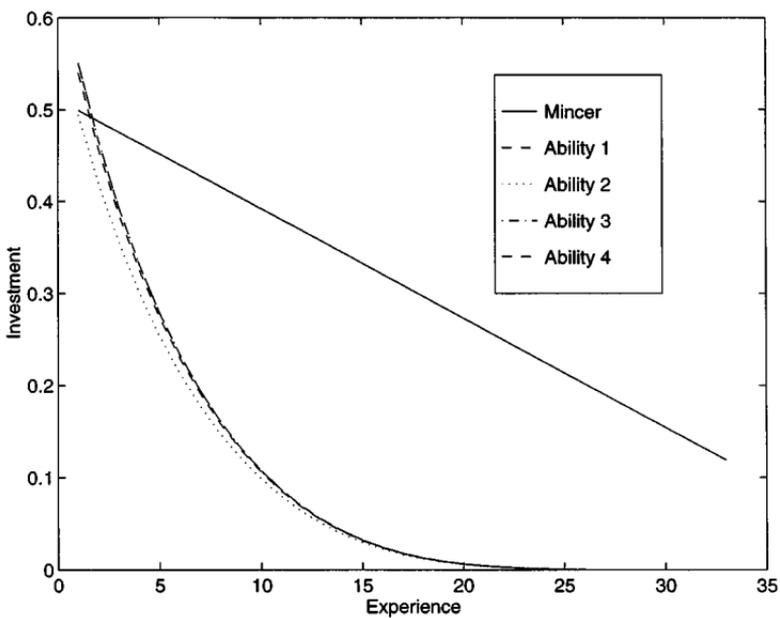
(b) college

FIG. 3.   Comparison of Mincer vs estimated investment profiles: (a) high school; (b) college.

θ can be written as

$$V^*_{(\theta)} = (1 - \tau)\left[V_i^2(\theta) - V_i^1(\theta)\right] - D_i + \varepsilon_i(\theta),$$

where τ is the tax rate and $\varepsilon_i(\theta) \sim N(\mu_\theta, \sigma_\varepsilon^2)$. We estimate μ_θ and σ_ε^2 by probit analysis using a two-step procedure. First, we run a probit regression of college attendance on tuition and dummy variables for each schooling–ability group for each of the seven birth cohorts in the NLSY. (The dummies estimate the difference in the after-tax valuation of schooling for each group.) The tuition variable is measured in units of thousands of dollars. We cannot reject the null hypothesis of no change in the value of attending school over the time period 1975–1982 when our sample makes its schooling decisions. For the sake of brevity, we report results (in Table I) for the case when the value of attending school is constrained to be the same across time.

The third column presents average derivatives. The interpretation of the average derivative is that an increase in tuition of \$1000 decreases the probability of attending college by about 0.08 on average. This estimate is on the high end of the range of estimates reported in the literature (see, e.g., [8] or [25]). In the second stage, we transform the parameters D_i presented in Table I into the structural parameters of our model. We use the relationships $\lambda = 1/\sigma_\varepsilon$ and $\alpha(\theta) = ((1 - \tau)[V^2(\theta) - V^1(\theta)] + \mu(\theta))/\sigma_\varepsilon$ and the estimates reported in Table I to form $[V^2(\theta) - V^1(\theta)](1 - \tau)$. Then we obtain

$$\mu(\theta) = \alpha(\theta)\sigma_\varepsilon - (1 - \tau)\left[V^2(\theta) - V^1(\theta)\right]$$

as the mean nonpecuniary return to college for a person of ability level θ. The estimates are reported in Table II. The only surprise in this table is the negative mean psychic cost of attending college for persons of the highest ability.

E. Estimating the Technology and Aggregate Stocks of Human Capital by Skill

Using the CPS data for the period 1963–1993, we employ the methodology presented in Sections II.A and II.B to estimate skill prices and human capital aggregates. The data sources for the macro aggregates are presented in Appendix A. From the constructed aggregates, we estimate aggregate technology (I.10) and test for various specifications of the aggregate technology. To correct for endogeneity of inputs, we use the standard instrumental variables often used in macroeconomics—military expenditures and cohort size. OLS and IV estimates of the technology are reported in Table III for all possible combinations of the instruments.

The estimated elasticity of substitution between capital and the labor aggregate Q is $\sigma_2 = 1/(1 - \rho_2)$ and is not statistically significantly different from 1. (See the fourth and fifth columns of Table III.) Our estimates justify excluding capital (or the interest rate) from Eq. (I.1). Our model produces rising wage inequality without assuming a special complementarity relationship between capital and skilled human capital—the centerpiece of the Krussell et al. [28] analysis of wage inequality and other discussions of rising wage inequality. Our estimates are consistent with the near-constancy of the capital share and the declining share of unskilled labor in the U.S. economy. (See Fig. 4.) The estimated elasticity of substitution between high-skill and low-skill labor (1.441) is remarkably close to the point estimates reported by Katz and Murphy (1.41) and Johnson (1.50). The instrumental variables estimators do not affect this estimate very much. Assuming no depreciation in human capital, we estimate the skill bias parameter as $\varphi = 0.036$, very close to the corresponding estimate of 0.033 reported by Katz and Murphy.

F. Calibrating the Model

Given our estimates of the human capital production function, we choose an initial steady state which is consistent with the assumptions used to obtain the estimates in the NLSY data. Given a tax rate τ ($= 0.15$) that was suggested by Pechman [37] as an accurate approximation to the true rate over our sample period once itemizations, deductions, and income-

TABLE III

Estimates of Aggregate Production Function (1965–1990) Estimated from Factor Demand Equations (II-1) and (II-4) Allowing for Technical Progress through a Linear Trend (standard errors in parentheses)

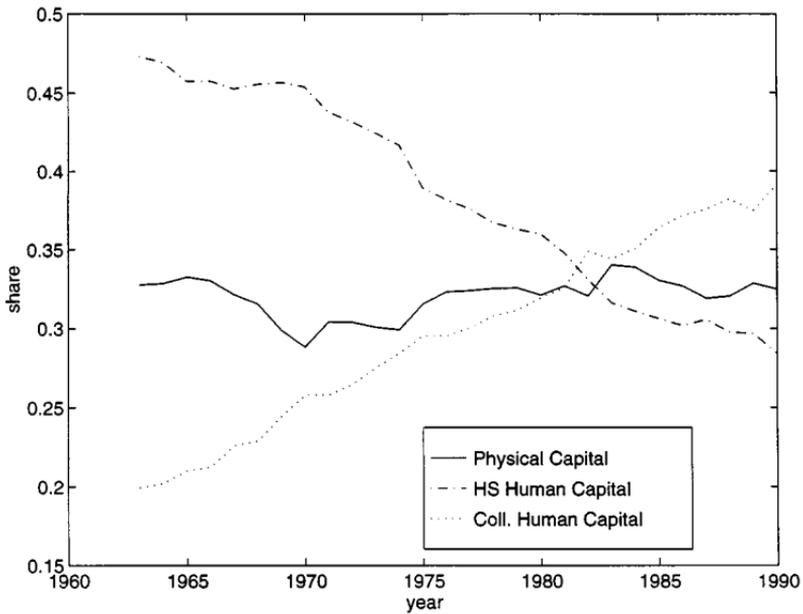| Instruments	(II-1)			(II-4)		
	ρ_1	Implied elasticity of substitution (σ_1)	Time trend	ρ_2	Implied elasticity of substitution (σ_2)	Time trend
OLS (base model)	0.306	1.441	0.036	-0.034	0.967	-0.004
	(0.089)	(0.185)	(0.004)	(0.200)	(0.187)	(0.007)
Percent working	0.209	1.264	0.039	-0.036	0.965	-0.004
pop. < 30 &	(0.134)	(0.215)	(0.005)	(0.200)	(0.187)	(0.007)
defense percent of GNP						
Defense percent	0.157	1.186	0.041	-0.171	0.854	-0.008
of GNP	(0.125)	(0.175)	(0.004)	(0.815)	(0.594)	(0.024)
Percent working	0.326	1.484	0.036	0.364	1.572	0.007
pop. < 30	(0.182)	(0.400)	(0.006)	(1.150)	(2.842)	(0.034)

FIG. 4. Labor and capital shares over time. Note: The breakdown of labor's share is based on wages and excludes other forms of compensation.

contingent benefits are factored in, we calibrate the aggregate production parameters to yield a steady state after-tax interest rate of 0.05 and pre-tax rental rates on human capital of 2. These values are consistent with those used in estimation of the human capital production parameters. Since human capital is measured in terms of hourly wages, earnings from our simulations are annual income measured in thousands of dollars if agents work 2000 hours per year. The calibration yields shares that are consistent with the NLSY white male sample we use.

It is necessary to make some assumptions about time preference and the intertemporal elasticity of substitution for consumption in order to determine savings rates and aggregate capital. Given the levels of human capital investment implied by the estimates of our model and levels of initial assets for each individual, we obtain consumption and savings under the assumption that $\delta = 0.96$ and $\gamma = 0.1$. In order to determine initial levels of assets, we partially redistribute physical capital from retiring workers (cohort a_R) to the cohort just entering the labor market so that the capital—output ratio in the economy is 4.[16] This calibration procedure yields an

---

[16] For each year, transfer X is taken from all workers at retirement age a_R, and the total amount is equally distributed to all individuals (irrespective of ability) of age 1 in that period. For the simulations reported in this paper, $X \approx \$30,000$.

initial steady state which emulates the NLSY data and central features of the macro economy. In Appendix C, we test the sensitivity of our simulations to the choice of these parameter values.

III. A DYNAMIC GENERAL EQUILIBRIUM MODEL OF RISING WAGE INEQUALITY IN THE U.S. ECONOMY

Using our estimated technology and human capital accumulation equations, we ask if we can explain the central features of rising wage inequality in the U.S. labor market over the period from the late 1970s to the early 1990s using a model of skill-biased technical change. The novel feature of our model, in contrast to the model of Bound and Johnson [5], Katz and Murphy [26], and Krussell et al. [28], is that we allow for endogenous skill formation. Ours is a model of a gradual shift in the skill bias of the technology to a higher permanent level.

Since we are using a human capital production function fit for young white males from the NLSY, our empirical model may not capture all features of the U.S. labor market. Since our estimates of the human capital production function for females are virtually identical to those of males, we do not think that this is a major source for concern.[17] Of potentially greater concern is our use of earnings data fitted on the early years of the life cycle for a recent cohort of workers. Since most human capital investment takes place early in the life cycle, we capture the main portion of such investment. Of more concern to us is the possibility of heterogeneity between cohorts in terms of endowments, ability, and human capital investment functions, which we ignore in this paper.

We consider a permanent shift in technology toward skilled labor using the estimate of trend parameter $\varphi = 0.036$ reported in Table III as our base case. We start from an initial steady state and suppose that the technology reported in Table III begins to manifest a skill bias in the mid 1970s. Greenwood and Yorukoglu [14] claim that 1974 was a watershed year for modern technology. Following their suggestion about the timing of the onset of technical change, $\log[(a_1)/(1 - a_1)]$ is assumed to decline linearly at 3.6% per year starting in the mid 1970s and continuing for 30 years. Shifts of longer and shorter duration produce qualitatively similar simulation results (within the time period we analyze) for a closely related base case we analyze in Appendix C.

To compute the general equilibrium of our model and the implied transition paths, we use the methodology of Auerbach and Kotlikoff [1, p. 213]. Starting from an initial steady state calibrated using the parameters

[17]See the estimates in [21].

of preferences and technology specified or estimated in Section II, we examine the transitional dynamics to the new steady state imposing the requirement that convergence occur in 200 periods or less.[18] Agents make their schooling and skill investment decisions under full rational expectations about future price paths, but they are surprised by a change in technology. Once the technology change begins, they know the entire future path of technology. For the same parameters and change in technology, a version of our model with myopic expectations does not converge to a new steady state solution but instead exhibits explosive cobweb behavior.

In the period immediately after the introduction of technology, the price of skilled human capital increases while that of unskilled human capital decreases. (See Fig. 5). This produces a rising college−high school wage differential (see Fig. 6), which is depressed initially, as more educated workers invest more and less educated workers invest less in on-the-job training. The ''rate of return'' in a Mincer regression increases 20−35% over the 10−15 year period after the technology shock begins. (See Fig. 7.) The return on physical capital declines slightly as the supply of human capital to the market declines, but then it rises as the human capital stock rises. (See Fig. 8.)[19] Movements in the return to capital are slight and play a small role in our story. An open-economy version of our model in which interest rates are set in world capital markets is qualitatively similar to the closed-economy story we tell here.

The skill price paths eventually stop accelerating as aggregate technology stops shifting and settles into its new steady state. Skill price differentials decline as the supply of skilled human capital continues to enter the economy until it reaches its new (and higher) steady state level.[20] The final rise in the price differential induces investment in skill to fit the requirements of the new technology. See Fig. 9 for the trend in the aggregate stock of potential and utilized skills produced from our model. Investment in human capital creates the wedge between the two measures of human capital stock.

An important feature of our description of the economy of the late 1970s and early 1980s is that the movement in wage differentials differs from the movement in price differentials. (See Fig. 10 where wage differ-

---

[18] Our models always converge in less than 200 periods so increasing the number of periods would not affect the transition path.

[19] The ''rate of return'' is the coefficient of schooling in a regression of log earnings on schooling, experience, and experience squared.

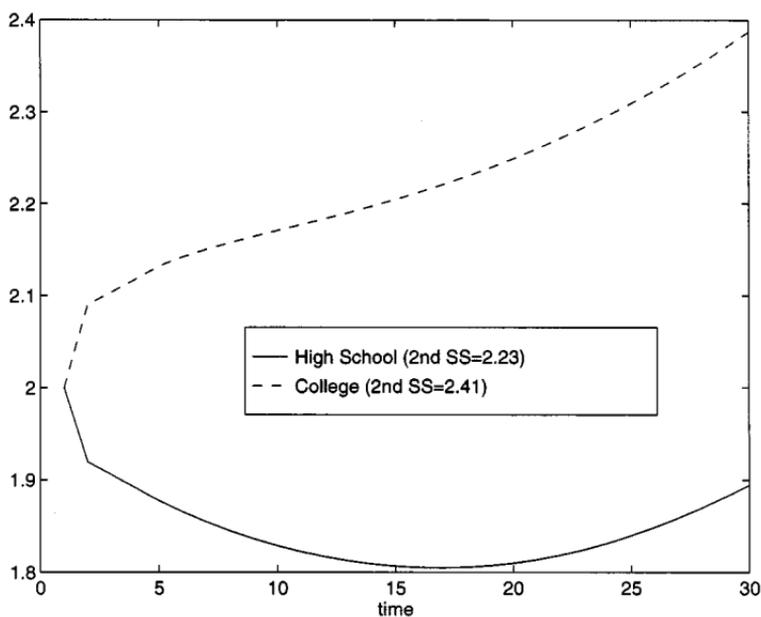[20] Graphs of the full transition are available on request from the authors.
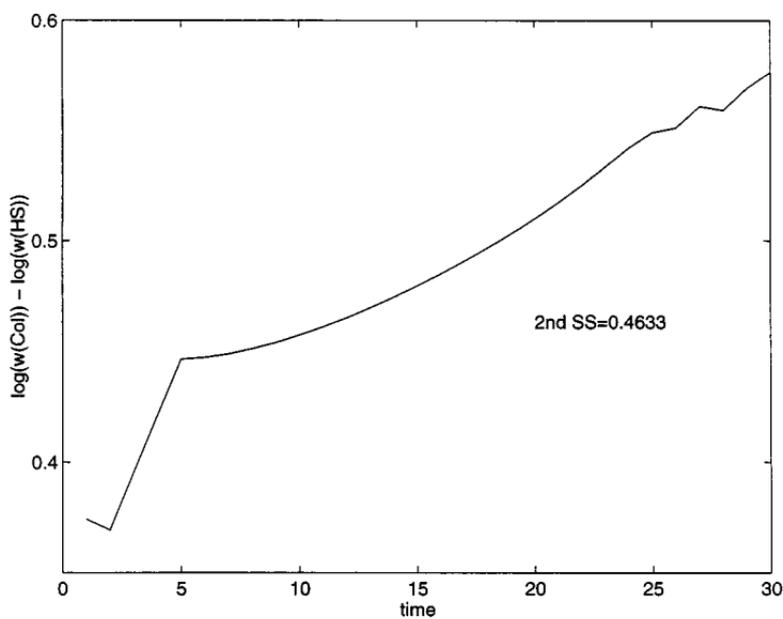
FIG. 5.   Prices of human capital.

FIG. 6.   College–high school wage differential.
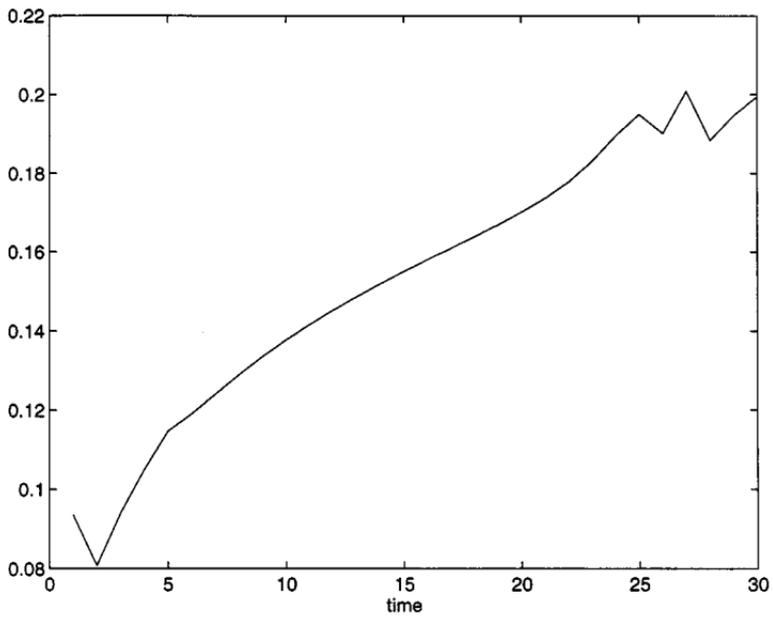
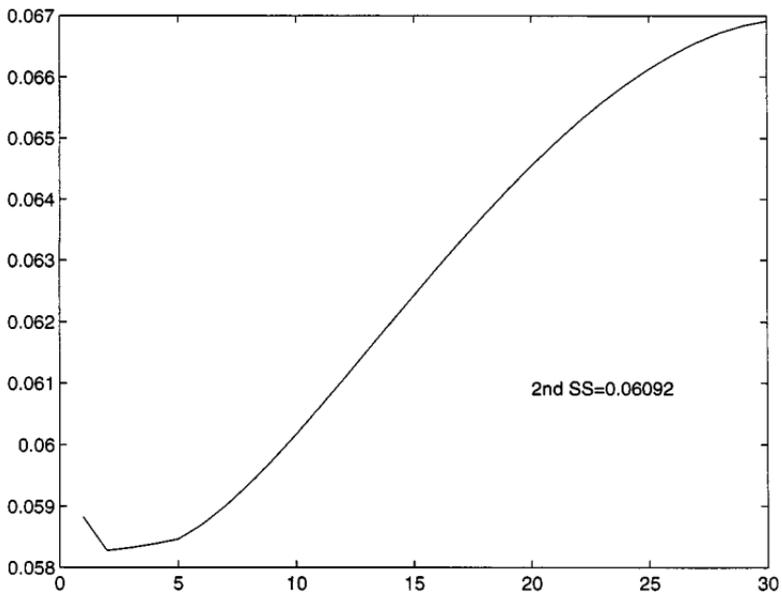**FIG. 7.** Schooling response in Mincer regression: "rate of return."
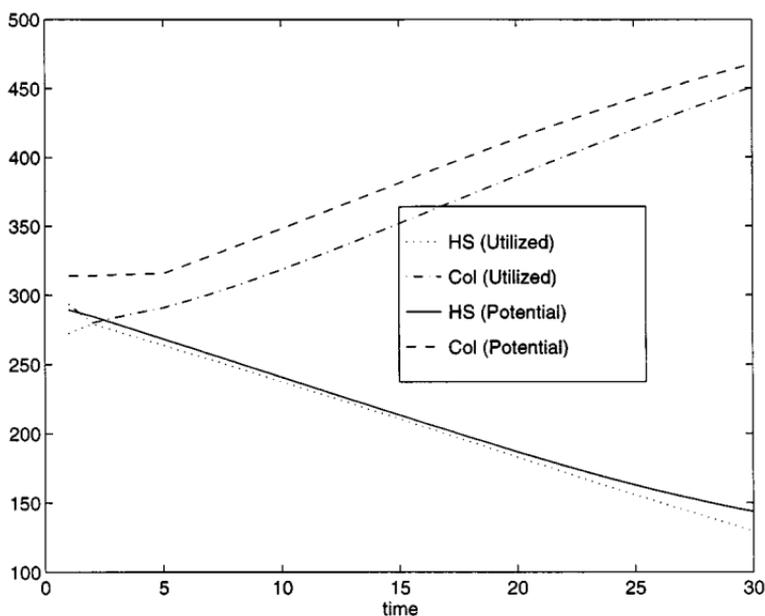

FIG. 8. Price of capital.

FIG. 9. Aggregate human capital used in production and aggregate human capital.

entials at different ages are compared to price differentials.) This phe-
nomenon is a consequence of the economics embodied in Eq. (I.11). In
response to the rise in skill prices in the years following the onset of
technological change, high-skill people who have left college invest *more*
on the job and then curtail their investment as opportunity costs of
investment increase relative to the payoff. These effects are especially
large for younger workers who are more active investors. Cohorts of young
skilled workers entering the market after the onset of technical change
invest substantially less than did earlier cohorts. (See Fig. 11a.) The
opposite is true for low-skill workers. For these workers, the price of skill
is initially high compared to where it ends up, so they invest much less in
the early years of the transition. Later cohorts invest more as the opportu-
nity costs of doing so decline.

This differential response in investment by skill groups over time ex-
plains the evidence for the 1980s presented in Katz and Murphy [26, Table
I] that the measured skill differential by education rises more for young
persons than it does for older persons.[21] Note that in the first phase of the
transition, differential investment by skill groups *narrows* the college–high
school wage gap, while in the second phase of transition, differences in

[21] This assumes that the onset of the technology shift is in the mid 1970s as claimed by
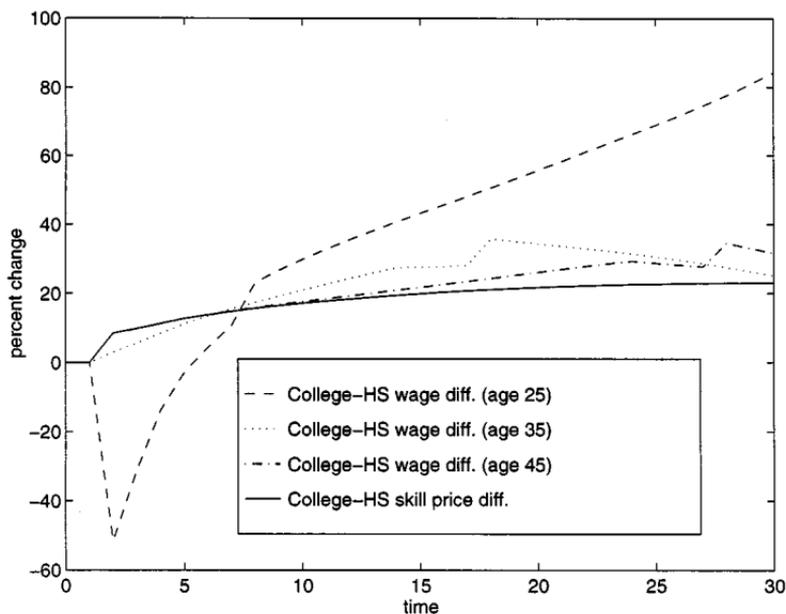Greenwood and Yorukoglu [14].

FIG. 10.   Percentage change (from initial SS) in wage rates and skill prices.

investment behavior widen the gap. In our model, the second phase of the transition has longer duration than the first phase. Our model in this phase is consistent with the evidence reported by Bartel and Sicherman [46]. They find that over the period at the mid-1980s, when wage inequality was increasing substantially, investment in company training increased for less educated workers both absolutely and compared to that for more educated workers in industries where technological progress was rapid.

In certain stages of the transition, movements at the intensive margin of skill investment are opposite to those at the extensive margin. Initially, both types of investment increase for high-skill workers, while on-the-job training declines for low-skill workers. However, in the second phase of the transition more people go to college, but both college- and high-school-educated workers invest less on the job for most age groups (as seen in Figs. 12a and b). In the long run, the amount of human capital per worker is *lower* for each skill type.

The model produces a large jump in college enrollments with the onset of technical change.[22] This is an artifact of our perfect foresight assump-

---

[22] For the sake of brevity we delete this figure. It is available on request from the authors.
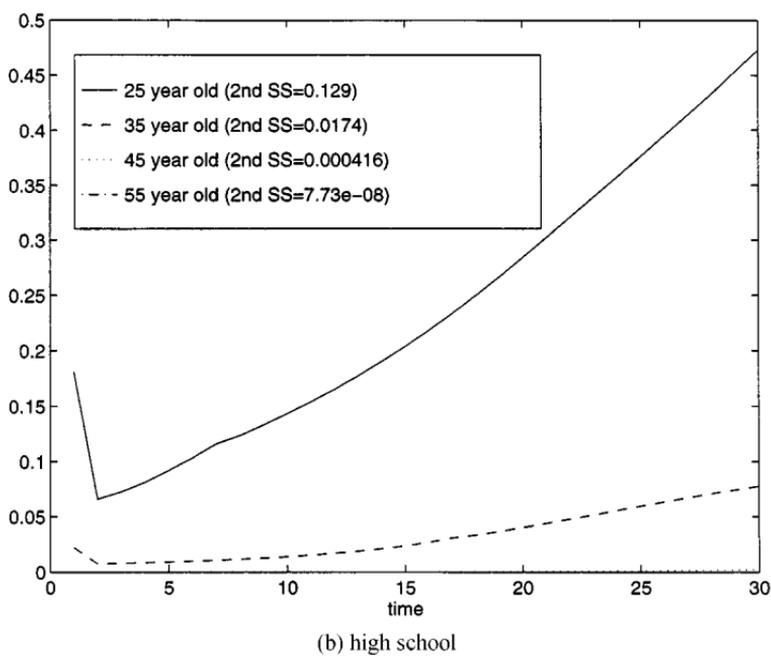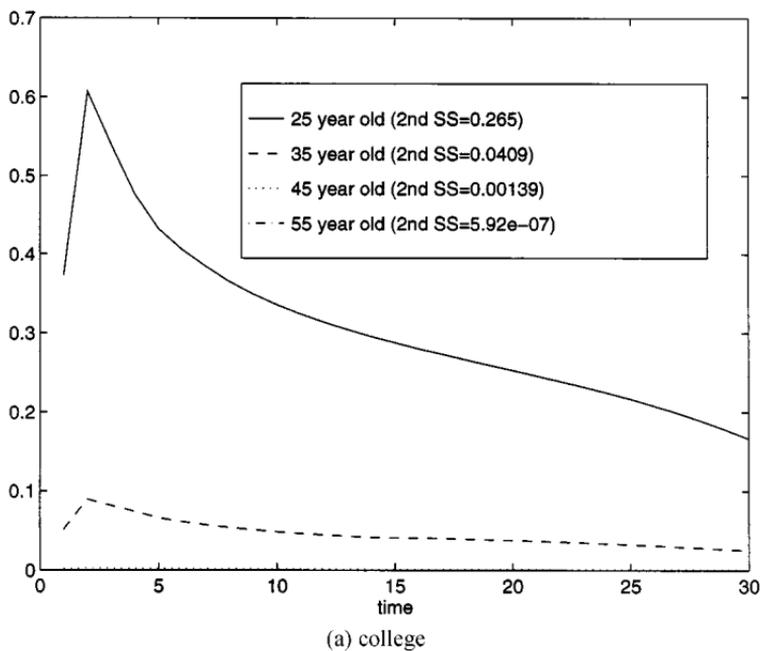
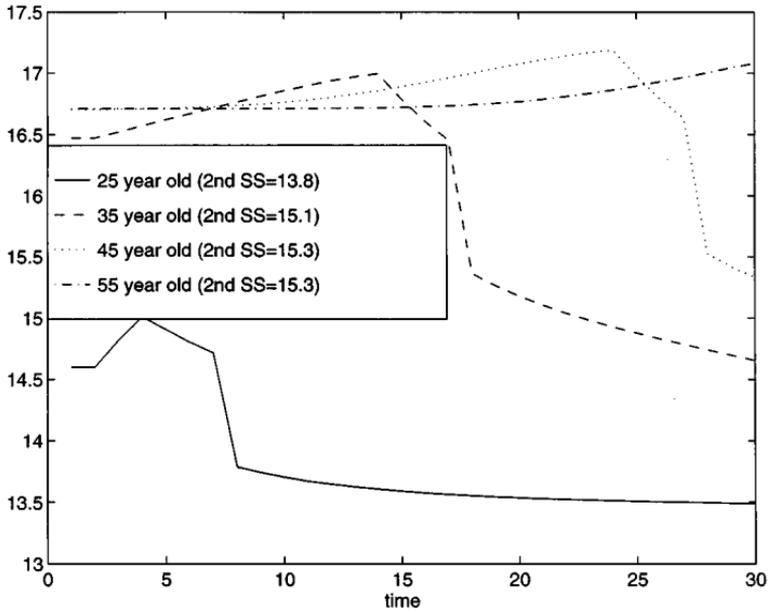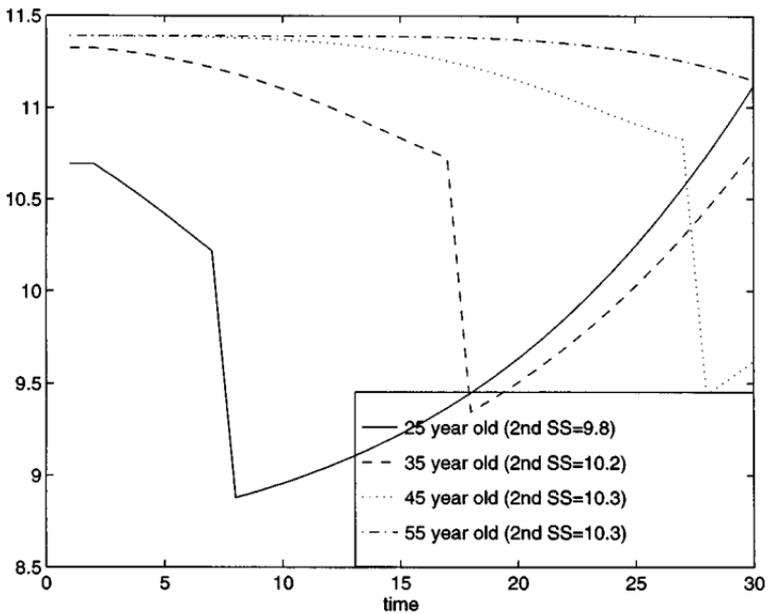(a) college

(b) high school

FIG. 11.   Proportion of time spent investing on the job: (a) college; (b) high school.

(a) college

(b) high school

FIG. 12. Average human capital: (a) college; (b) high school.

tion. A model in which information about skill bias disseminates more slowly would be more concordant with the data.[23]

It is interesting to examine the self-correcting properties of this equilibrium. In response to the new technology regime, the standard deviation of log wages initially rises sharply but then converges to a *lower* steady state value. (See Fig. 13.) The model also explains rising wage inequality at different percentiles of the wage distribution. However, this phenomenon is transient. (See Fig. 14.) After the initial phase, inequality increases within the low-skilled group as measured by the variance of log wages. Initially, the variance within high-skill groups increases, but eventually it declines. On the latter point, our model is at odds with the stylized facts about wage inequality.[24]

The models of Krussell et al. [28] and Greenwood and Yorukoglu [14] abstract from heterogeneity within skill groups and cohorts and also cannot explain the within-skill-class rise in inequality. Their models cannot explain the flattening of wage−experience profiles for high-skill groups or the steepening of wage experience profiles for low-skill groups. The analyses of Caselli [10] and Violante [44] also cannot explain this phenomenon. For Caselli, individuals work only one period, so his model offers no prediction about experience−wage profiles. Violante's model predicts that when a new technology arrives, the returns to skills learned on previous technologies fall, leading to a *decline* in the slope of wage−experience profiles for low-skill workers. This prediction is grossly at odds with the data. However, both Caselli's and Violante's models explain the rise in wage inequality within narrowly defined skill cells. For Violante, this is a consequence of matching of workers to vintages, accelerated technical change, and induced labor turnover. For Caselli, this is due to increased sorting of high-skill labor with capital, where skill is endogeneously determined.

Appendix C presents a sensitivity analysis that examines the robustness of the simulation estimates reported here to alternative assumptions about

[23]An alternative way to get a more gradual response in educational enrollment is to endogenize tuition, recognizing that college is skilled-labor intensive so that as skill prices increase, the cost of schooling rises. A two-sector model could substantially dampen the jump in estimated college enrollment. If we assume myopic expectations on the part of agents, the jump in enrollment in college at the onset of technological change is *larger*.

[24]Our model is consistent with the more recent evidence. Krueger (personal communication, 1997) suggests that rising wage inequality within narrowly defined skill groups is no longer increasing for all skill groups. When we alter the model to account for migration into the economy of low ability workers (from the lowest quartile) we can produce widening wage inequality within all skill groups. However the required increase in migration is implausibly large.
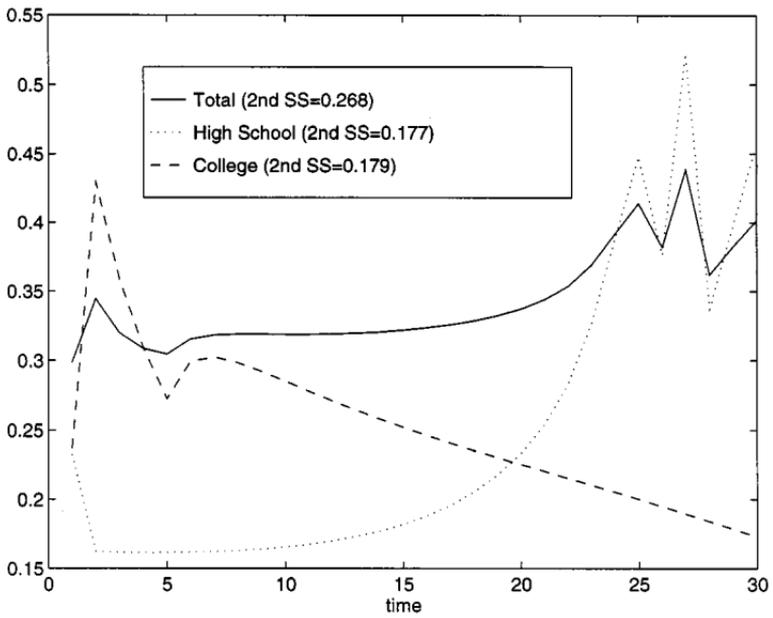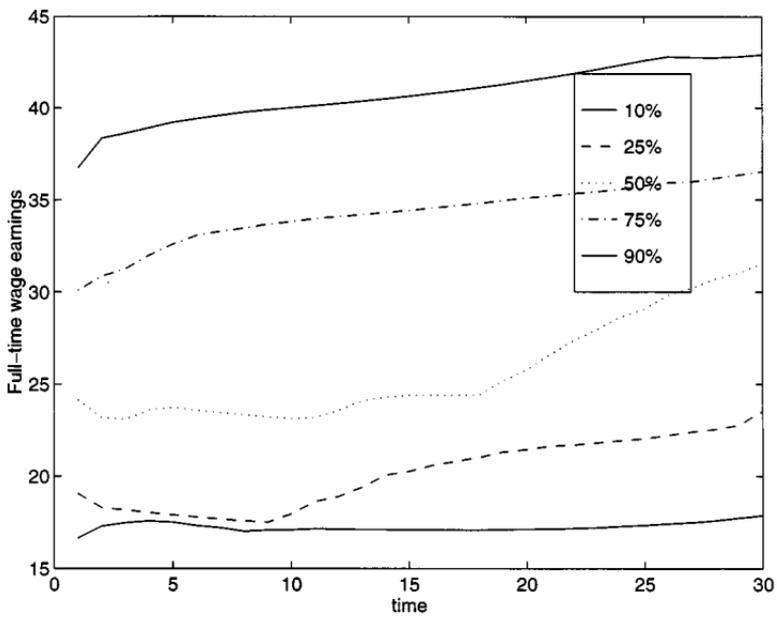
FIG. 13. Log wages—standard deviation.

**FIG. 14.** Percentiles of wage distribution.

model parameters. Qualitatively, our main conclusions are robust. Quantitatively, different choices of parameters affect the magnitudes of the simulation outputs.

## A. Comparisons between Cohorts and Cross Sections

One benefit of an overlapping generations model is that it enables us to compare and contrast cross sections with cohort paths that are economically more interpretable for conducting welfare analysis. Our model enables us to engage in systematic generational accounting in the fashion pioneered by Auerbach and Kotlikoff [1]. Figures 15a−15d informs us how the overall lifetime utility, the lifetime utility by ability, and the lifetime utility by ability and schooling type change over cohorts.

The widely used Benthamite measure of aggregate lifetime utility (obtained by summing over the utilities of all persons in the economy at a point in time) declines for cohorts entering the labor market just prior to the onset of the technological change. (See Fig. 15a.) The date of onset of the technology change is at the time cohort zero enters the labor market. Utility increases for cohorts entering the labor market after the onset of technical change, but successor cohorts have less utility than the initial cohorts. Later cohorts gain even larger utility. Disaggregating by ability groups (Fig. 15b), higher ability workers gain from the new economy both in the long run and in the short run. High ability persons who enter the economy before the technology shock do somewhat better than their predecessors. Cohorts of lower ability persons entering the economy before the technology shock do worse. Low ability cohorts born before the onset of the technology are hurt for substantial periods of time after the shock occurs. In the long run, cohorts of workers of all ability levels are better off.

Figures 15c and d report results disaggregating by ability and education. The utility path for high school-educated workers declines but recovers for cohorts born after the shock until it reaches a new higher level. (See Fig. 15c.) Cohorts born immediately after the shock have lower utility than the predecessor and successor cohorts. In the new steady state, cohort utility levels are higher. For college-educated persons, the story is different and is not entirely the mirror image of the case for high school graduates. (See Fig. 15d.) High-skill persons educated before the advent of technology change capture a large rent due to the unanticipated rise in skill prices. Successor cohorts do not fare as well. Their lower mean utility can be attributed to the strong distaste for college of the new persons attracted into college. They now attend college because the decline in their earnings in the unskilled sector is even greater than the tuition and psychic costs of
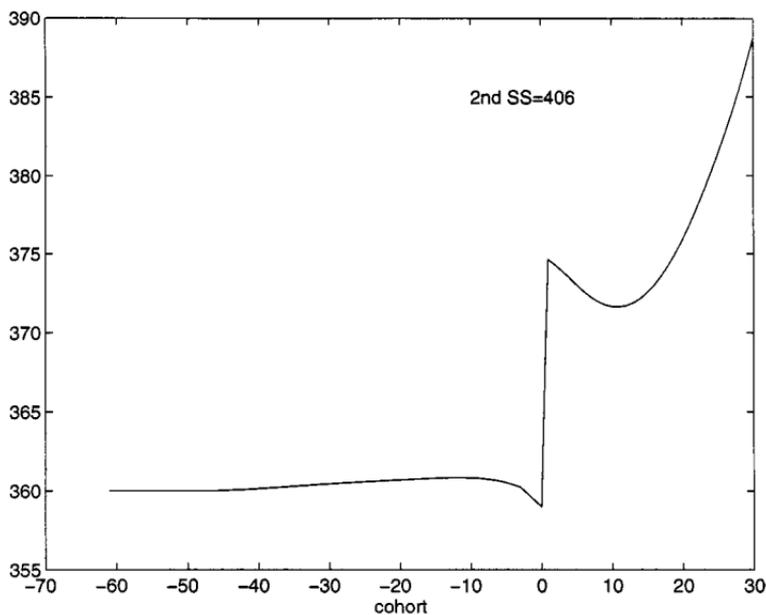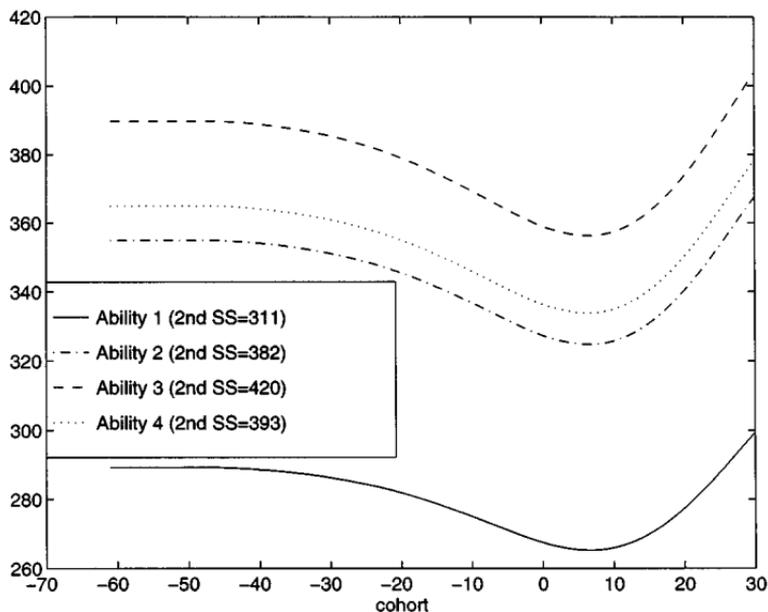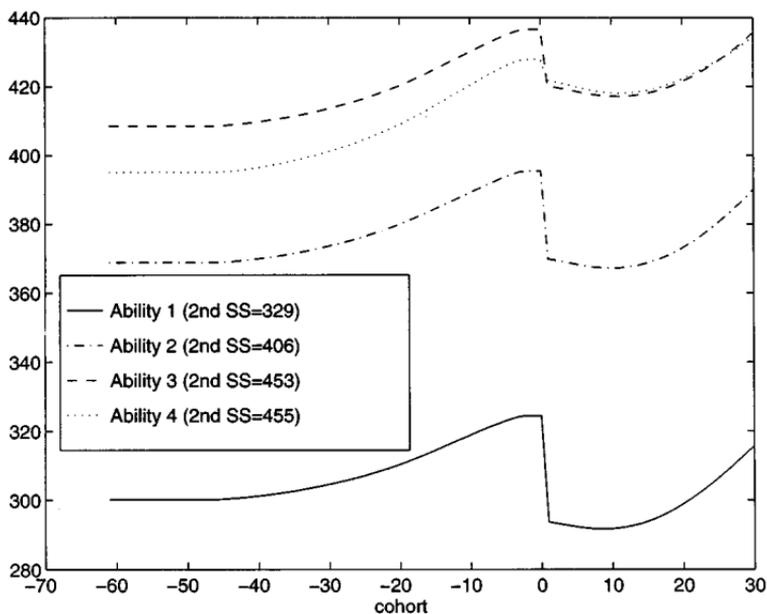
(a) Weighted average lifetime utility
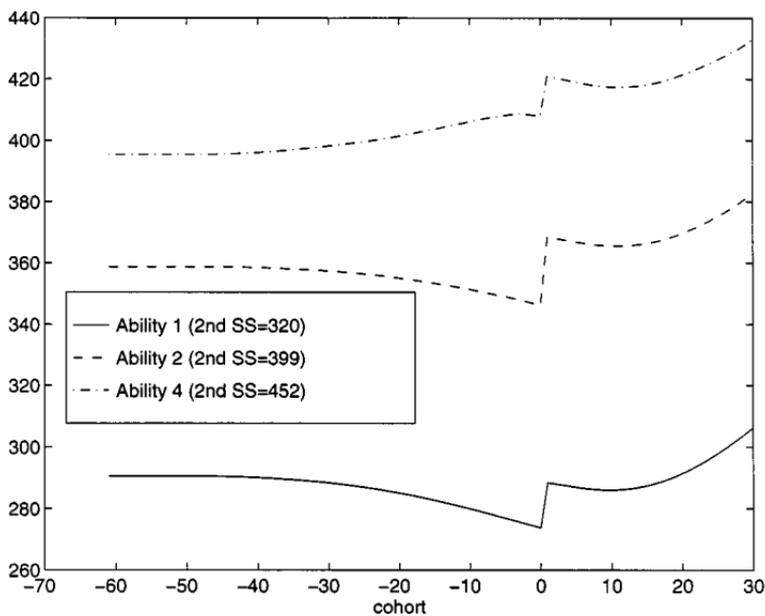
(b) Weighted average lifetime utility by ability

FIG. 15. (a) Weighted average lifetime utility; (b) weighted average lifetime utility by ability; (c) lifetime utility—high school graduates; (d) lifetime utility—college graduates.

(c) Lifetime utility—high school graduates

(d) Lifetime utility—college graduates

FIGURE 15—*Continued*

attending college.[25] For people who are induced to attend college, schooling is now better than the new low-skill alternative, but they are worse off than they would have been if had they remained high school graduates in a previous era. This group lowers the mean utility of all college graduates.

The pattern of declining lifetime utility for cohorts of both less-educated and more-educated workers a decade or so after the onset of technological change is consistent with the evidence reported by MaCurdy and Mroz [32] and Beaudry and Green [3] that cohorts entering the labor market immediately after the start of technological change do worse than predecessor cohorts. (This pattern is also true for the net present value of labor earnings. See Fig. 16.) Our model predicts that this pattern will be reversed in the long run as the forces of technology attenuate and as supplies adjust to eliminate wage differentials.

The important role of nonpecuniary costs in explaining college attendance accounts, in part, for the gap between the opportunity cost of capital (5%) and the "Mincer return," which ranges from 9.5 to 14% (see Fig. 7). Nonpecuniary components are 15% of the total return to college atten-

[25] The discontinuity in the utility paths for college graduates arises from the discontinuity in college attendance induced by the onset of technology change and by the greater psychic and tuition costs of attending school by the marginal college entrants.
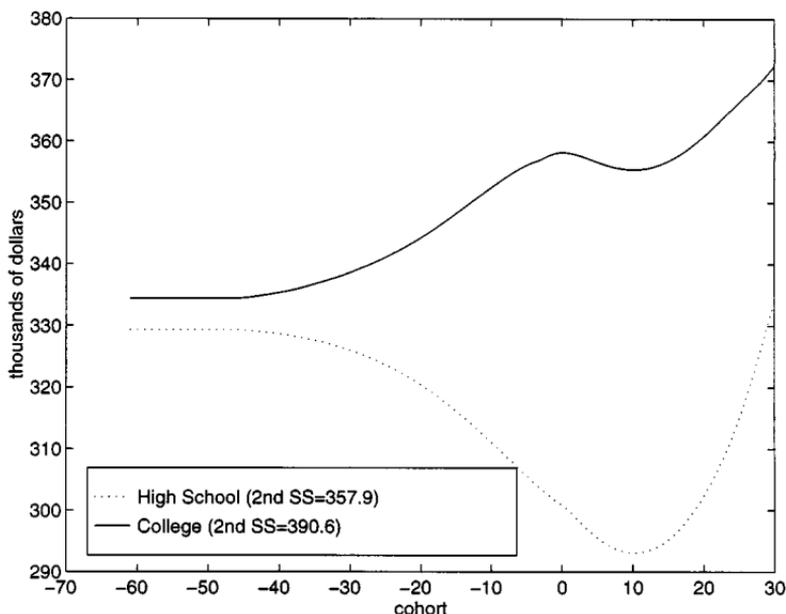

FIG. 16. Discounted lifetime earnings.

dance for the more able. Also, observe the gap between the cohort "rate of return" (Fig. 17) and the cross-section rate (Fig. 7). During the transition, the estimated cross-section "rate of return" is one to two percentage points higher than the "rate of return" experienced by any entry cohort. Cross-sectional "rates of return" are not appropriate guides to educational investments for entering cohorts, although they are often used that way. This points out an important weakness in the conventional method of evaluating tuition and other policies in addition to the more familiar problems that (a) monetary rates of return are not true rates of return (because of psychic benefits) and (b) that steady state general equilibrium adjustments resulting from policies are typically ignored. During periods of transition to a new skill regime, cross-section rates of return to education present an overly optimistic account of the "rate of return" any single cohort can earn. For example, the Mincer coefficient in year 15 is around 15% (Fig. 7) while the Mincer return for the cohort entering in year 15 is 11.5% (see Fig. 17).[26]

[26] Heckman and Klenow [20] decompose the Mincer coefficient into components due to (a) tuition costs, (b) nonseparability between schooling and work experience, (c) the returns to schooling due to longevity extension, and (d) uncertainty. The Mincer coefficient is a "true" rate of return that can be compared to the interest rate only under very special circumstances.
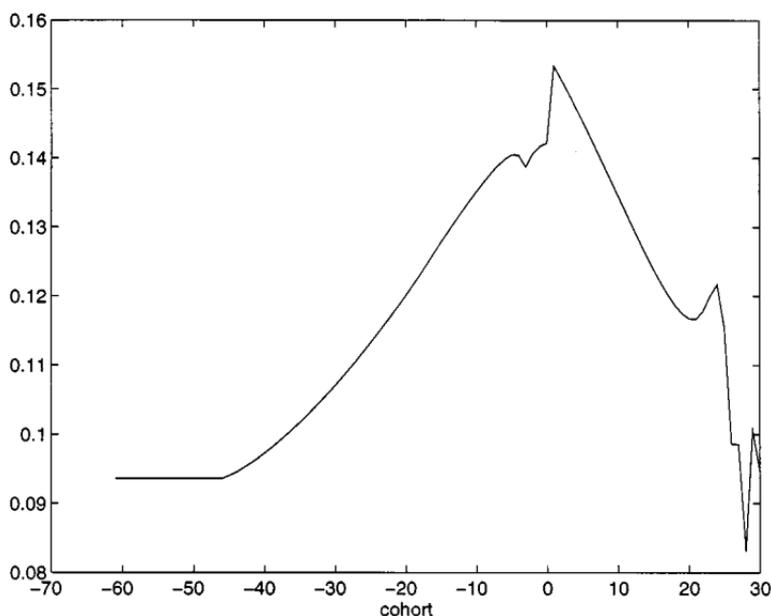

FIG. 17.   Schooling response in Mincer regression: "rate of return."

Observe, finally, that within-cohort wage inequality, presented in Fig. 18, often moves opposite to aggregate wage inequality. After the technology shock, entry level cohorts experience less aggregate wage inequality than predecessor cohorts even though aggregate wage inequality is rising (compare Figs. 14 and 18). In year 20 of the technology change, the 90−10 differential is growing compared to previous years. (See Fig. 14). For cohort 20, the 90−10 differential narrows compared to that of predecessor cohorts.

B. The Effect of Immigration of Low Ability People on Wage Differentials

Immigration of unskilled workers is sometimes considered to be a rival explanation to skill-biased technical change in accounting for rising wage inequality. In order to examine the quantitative importance of this explanation, we simulate migration and expand the supply of low ability (θ) persons into the economy over a 50 year period. In order to capture the relatively low skills of immigrants, we consider cases where they are dumped into the economy past the schooling age and are the lowest ability among the "high school" graduates. A variety of simulations, deleted for the sake of brevity, confirm that a 25% increase in migration of the unskilled operates to widen the college−high school skill differential, and increases the standard deviation of log earnings within both skill groups.
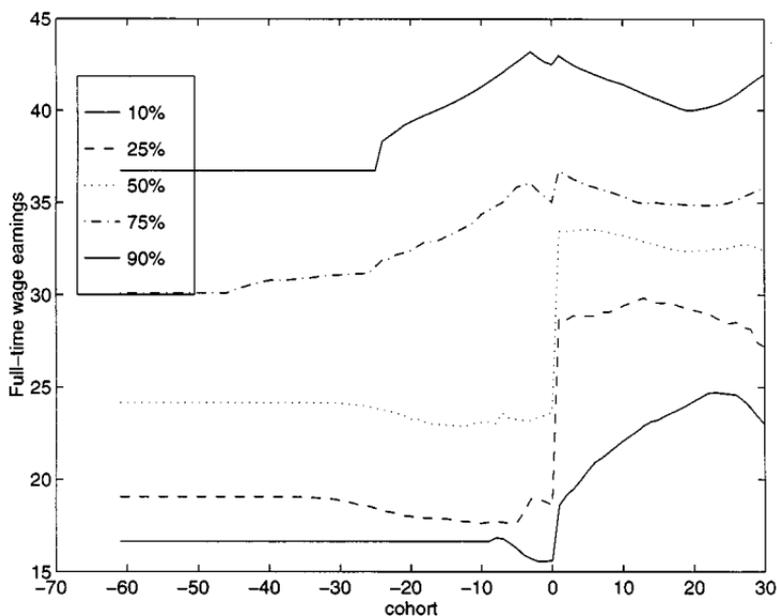
FIG. 18.   Percentiles of wage distribution.

However, its effect is too small to account for the rise of 10–11 log points in the college–high school wage differential.[27] At most we can explain 0.01 log points. Implausibly large increases in migration are required to account for the growth in wage differentials witnessed in the past 15 years.

C. Accounting for the Baby Boom

As a test of our model, we ask if we can account for the past 35 years of U.S. wage history using the aggregate technology and human capital accumulation equations estimated in this paper. We consider an economy in which an episode of skilled-biased technical change ($\varphi = 0.036$) begins in 1960 and continues for 30–40 years. To this we add the demography of the Baby Boom in which cohort sizes increased by approximately 32%. We assume that Baby Boom cohorts begin to enter the economy in the mid 1960s and continue for a period of 15 years.

Figure 19 presents the simulated college–high school wage differential. It captures the essential features of recent U.S. wage history. The differential increases in the 1960s, decreases in the 1970s, and rises again in the 1980s and 1990s. In results not shown here, the simulated model predicts

[27] These simulations are available on request from the authors.
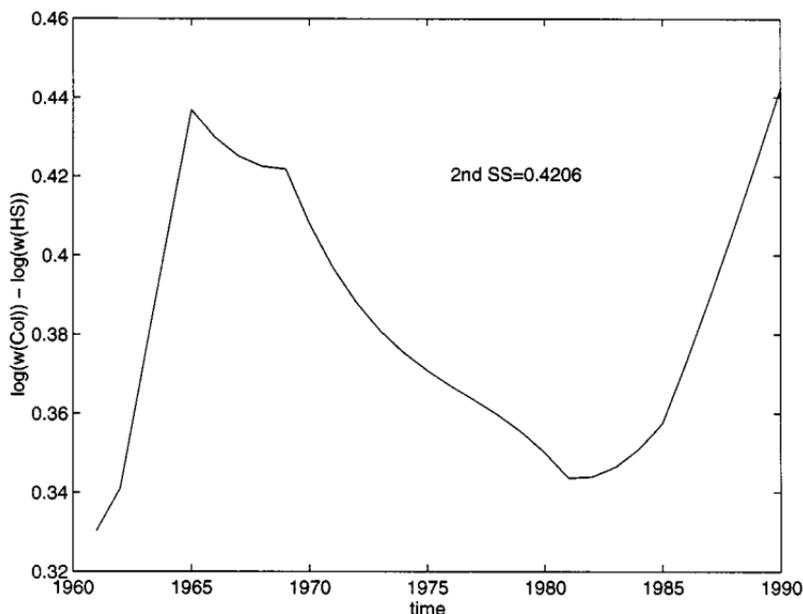
FIG. 19. Baby Boom (expansion of cohort size by 32%) between years 1965 and 1980. College–high school wage differential.

that college enrollment rates rise in the 1960s, decline in the 1970s, and rise again in the 1980s, and real wages of high school graduates rise in the 1960s and 1970s, but decline in the 1980s and 1990s. The declines in the wages of high school graduates are more moderate in the 1980s than they are in the base case previously analyzed. The real wages of college graduates rise, fall, and then rise in the period of the 1960s, 1970s, and 1980s, respectively. The college–high school wage differential at age 25 rises over the period 1963–1971, decreases until the early 1980s, and then rises again. (See Fig. 20.) Overall, inequality measured by the standard deviation of log wages rises starting in the mid to late 1960s and generally increases, taking a slight downward turn in the late 1970s and early 1980s before rebounding with renewed vigor in the 1980s.

Table IV summarizes the properties of our model in a format comparable to Table I of [26]. The college–high school wage differential rises in the 1960s, falls in the 1970s and rises in the 1980s in a fashion that mirrors the evidence reported by Katz and Murphy [26, Table I]. Income inequality measured by the standard deviation in log wages increases over all decades, although the largest change is in the 1980s.

With the same basic ingredients of human capital investment and aggregate technology, we can explain 35 years of U.S. wage history,

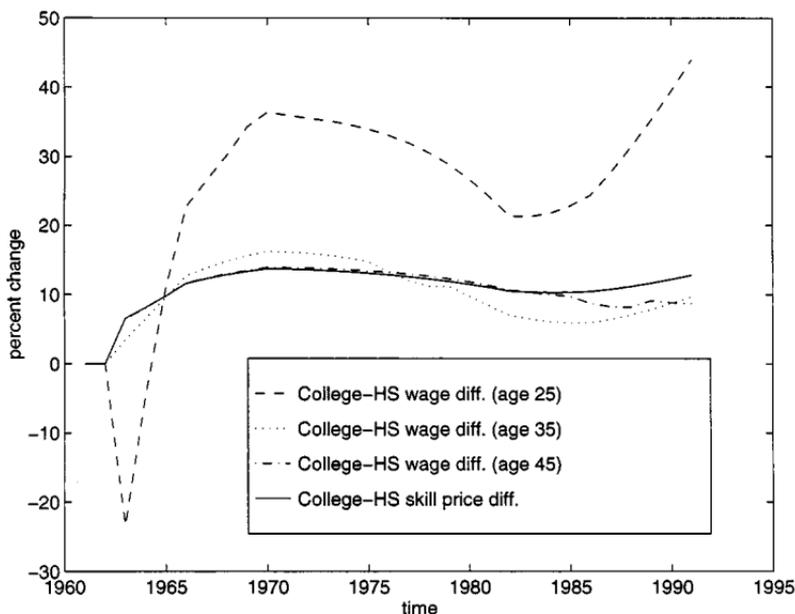**FIG. 20.** Baby Boom (expansion of cohort size by 32%) between years 1965 and 1980. Percentage change (from initial SS) in wage rates and skill prices.

TABLE IV

Simulated Percentage Changes in Wages and Wage Inequality from 1960 to 1990 Including
the Estimated Trend in Technology and Entrance of Baby Boom Cohorts from
1965 to 1980

Years	Coll.−HS log wage diff.	Mean HS log wage		Mean coll. log wage		Std deviation of log wages		
		Age 25	Age 50	Age 25	Age 50	HS	College	All
1960−1970	6.66	−26.98	−9.17	19.41	−2.2	0.06	0.67	2.49
1970−1980	−5.33	3.51	−2.32	−8.72	−5.11	2.06	−0.84	0.14
1980−1990	11.74	−4.94	−1.74	11.22	−2.72	10.68	−7.87	8.12
1960−1990	13.07	−28.4	−13.22	21.91	−10.03	12.8	−8.03	10.75

assuming that skill-biased technology starts around 1960 and continues for 30−40 years and the Baby Boom cohorts enter the work and schooling economy in the mid 1960s. The expansion in college enrollment in the 1960s can be explained by basic economic forces and not as consequence of generous tuition policies.

IV. SUMMARY AND CONCLUSIONS

This paper develops an empirically grounded dynamic overlapping-generations general equilibrium model of skill formation with heterogenous human capital that explains rising wage inequality as a consequence of skill-biased technical change. Our model extends the pioneering model of Ben Porath [4] to a market setting by relaxing his assumption of efficiency units for labor services. Instead we consider a model with comparative advantage in the labor market. Schooling human capital is distinguished from human capital acquired on the job. Human capital investment choices are considered both at the extensive margin (schooling) and at the intensive margin (OJT). We allow for heterogeneity in ability and produce a model that is consistent with the main features of life-cycle wage growth and the growth in wage inequality that are central features of the recent U.S. labor market. Our model also explains why cohorts that enter the labor market after a technology shock occurs are worse off, at least in the short run.

Distinguishing between skill prices and wages is important for interpreting changes in the wage structure. The link is broken by human capital investment decisions which themselves depend on price paths. Once it is recognized that wages are not prices, it is possible to understand the recent episode of rising wage inequality. In later stages of the transition to a new

level of skill bias in technology, the rise in the college−high school wage differential overstates the rise in skill prices as a consequence of rational investment behavior, especially for young workers. As a result, college wage profiles flatten while high school profiles steepen during what we identify as the second phase of the transition. In the first phase of the transition, we find that measured college−high school wage differentials *narrow*. This highlights the central point in our paper that wage differentials are poor signals of skill price differentials. In the first phase of the transition, wages and prices move in *opposite* directions.

We present new methods for measuring unobserved aggregate stocks of skill-specific human capital. These methods are of interest in their own right, because they provide us with the ingredients to conduct more accurate growth accounting that allows for unmeasured on-the-job investments. Using these methods, we estimate the aggregate technology linking output to capital and human capital, the elasticity of substitution between skilled labor and unskilled labor, and the elasticity of substitution between the skill aggregate and capital. We find that capital−skill complementarity is not required to produce rising wage inequality.

We use our model to examine the quantitative importance of the migration of unskilled workers to aggregate wage inequality. For plausible increases in the pool of unskilled labor, we find that migration cannot account for much of the growth in wage inequality.

We test our framework by building a model with a Baby Boom in entry cohorts and an increase in the bias of aggregate technology that starts in the early 1960s. With these ingredients, we are able to produce a model consistent with the central facts of U.S. wage inequality over the past 35 years.

APPENDIX A

Data

The data used to estimate the human capital production function are from the National Longitudinal Survey of Youth (NLSY)—a nationally representative sample of individuals that began in 1979 and interviewed youth aged 14−22. These same individuals were reinterviewed annually until 1993—the last year of data that we use. We use a subsample of white male civilians from the NLSY and exclude the oversampling of poor whites. Individuals are included in the sample if they work more than 500 hours in a particular year. Persons with hourly wages above $100.00 and less than $1.00 in 1992 dollars are deleted.

We partition the data into four groups on the basis of their Armed Forces Qualifying Test (AFQT) score. These identify the θ types as used

in this paper. In 1980, 94% of the sample was given the Armed Services Vocational Aptitude Test, which consists of 10 standardized tests that are used to assess a variety of skills. Four of these tests are combined to form the AFQT, which is used as an admission criterion into the armed forces. We normalize the test by subtracting the mean score for each individual's birth year. We then divide the subsample of white males into four equal-sized groups ranked on the basis of their AFQT score.

In estimating the price elasticity of the decision to attend college, we use the state average tuition levels for public two year institutions (from the calendar year in which the sample member turned 18) as our measure of the local price of college.

The macro aggregates come from the National Income and Product accounts as presented in the Citibase data and in FRB data sets on capital stock. We define labor's share in the following way:

labor's share = compensation/

$$(GNP - \text{indirect business taxes} - \text{proprietors' income}),$$

capital's share = 1 − labor's share.

Note that indirect business taxes and proprietors' income are, equivalently, being excluded from the calculation and are assumed to break down the same way that the rest of GNP does between labor and capital. Indirect business taxes are largely sales taxes, which are "skimmed" off before businesses can allocate the income to capital or labor, and proprietors' income includes income of the self-employed and nonincorporated partnerships (family businesses, law firms, etc). We do not know how to break down proprietor's income, because it is compensation to both labor and capital.

APPENDIX B

Sensitivity Analysis

In this appendix we analyze the sensitivity of our results to estimates produced from various widely used strategies. We perform the following simulations using samples constructed from our artificial economies. (1) From simulated economies, we generate Monte Carlo samples of micro data on wages, taking random draws from the distribution of schooling and from the distribution of measurement error. (2) We then estimate the micro parameters from this Monte Carlo data, misspecifying features of the economic environment using misspecifications that are common in the literature. (3) Finally, using the estimates of the misspecified models we

recalibrate the general equilibrium model to achieve desired capital–output, share parameters, and initial prices. We then see how well we predict the rise in wage inequality examined in this paper. We consider four separate cases as outlined in Table B-1.

Since our goal is to test the sensitivity of the micro-estimation-imposed methods to alternative misspecification of the economic environment, we minimize the role of sampling variance by choosing the standard deviation of the measurement error to be small (1/10 of the estimated standard deviation).[28] We obtain predicted probabilities of college for each cohort and each group, and we draw random samples from this distribution of probabilities. In the first benchmark experiment, we estimate the model under correct assumptions about how the micro data are generated. We take micro data from the original steady state and then estimate the parameters of the model using the correct assumptions about the way the data are generated. As expected, the simulated data in this case look very similar to the original model.

To simulate estimates of the data with changing prices, (experiment (2)) we use data from the economy undergoing change in the share of unskilled labor reported in the text. To approximate the NLSY, we choose eight cohorts and assume that we have data on each cohort from entry into the labor force until age 32. The introduction of technical change comes in the middle of these cohorts. We have 15 observations on wages for high school graduates and 11 observations for college graduates. We choose these eight cohorts to give us substantial price variation over the constructed sample period. We take four cohorts that make college decisions immediately preceding the technology change and four cohorts that make their enrollment decision immediately following it.

We find that the base case 30 year technology change simulated using these estimated parameters is very close to our base simulations. The

---

[28] For the schooling parameters there is no measurement error, so there is no analogous normalization.

TABLE B-1
Models Investigated in the Sensitivity Analysis

Problem studied	The true economic environment	Assumptions made in estimation	Assumptions made in calibration
(1) Benchmark run	Steady state	Steady state	Steady state
(2) Misspecify skill prices	Transition economy	Steady state	Steady state
(3) Misspecify interest rate	$r = 0.05$	$r = 0.10$	$r = 0.05$
(4) Misspecify interest rate 2	$r = 0.05$	$r = 0.10$	$r = 0.10$

magnitudes of some of the effects are somewhat different, but all of the qualitative features of our base simulations appear in these simulations as well.

The third and fourth experiments involve estimating the model under an incorrect assumption about the interest rate. We use the simulated data set from the benchmark run, but in estimating the parameters of the human capital production function, we impose the incorrect assumption that the interest rate is 10%. We then calibrate the general model using two alternative assumptions. In the first we use the estimated parameters and the correct macro interest rate of 5%.[29] In the second, we use the same estimates but calibrate to an intitial steady state interest rate of 10%. Thus, in the first case, the true model and the simulated model correspond, but the estimated model is different. In the second case, the estimated model and simulated model correspond, but they are different from the true model.

Our third case yields very poor results. When we estimate the model using a 10% interest rate but calibrate with a 5% rate in the original steady state, the simulations do poorly. The problem is that cutting the interest rate in half substantially increases the amount of on-the-job training. In these simulations, we find that individuals spend all of their work time in on-the-job training until about age 45. Examining the effects of techonology change is extremely difficult and leads to uninterpretable results. In the fourth case, where we estimate and simulate the model using the same interest rate, the model performs much better. As in the second case, we obtain the same qualitative results in response to changes in technology as in the base case. It is clearly important that the estimated and simulated interest rates correspond.

APPENDIX C

Sensitivity to Parameter Estimates, Cohort Distribution, and Duration of Technology Change

We perform a sensitivity analysis of our model to determine whether changes in estimated and calibrated parameters have any effects on our main findings. We also examine whether changes in the distribution of various cohorts affect our results. Finally, we examine the effects of varying the duration of technology change. Nearly all of these changes result in slight quantitative differences; however, the same general story about wage inequality holds.

---

[29] In obtaining estimates of the parameters of the schooling equation we use present values that are constructed using a 5% interest rate.

In Tables C-1 to C-3, we show the effects of each variant of our model on wage inequality and wages over the first 10 years (Table C-1), the first 15 years (Table C-2), or first 20 years (Table C-3), of the transition period. These years are intended to represent the late 1970s onward. It is important to note that the base case in this appendix differs from the base case presented in the main paper. (This sensitivity analysis was done for an earlier version of this paper.) The only difference between this base case and the one in the paper is the rate of decline in technology parameter a_1. In this appendix, the base model is chosen to yield a linear decline in a_1 over 30 years such that the share of low-skill human capital in the economy declines a total of 30% from its initial steady state level. This was the model we initially explored, but we later abandoned it in favor of the case reported in the text of the paper. Our old base case explains too little of the rise in the college−high school wage differential. Otherwise the qualitative properties of both models are the same. (In the base case of the main paper, the rate of decline in a_1 is taken from our estimates of the aggregate production function and yields a total decline of about 65% in low-skill human capital over 30 years.)

As seen in Table C-1, the base case in these simulations produces a smaller rise in the college−high school wage differential than is produced from the main model estimated in the text (8.67 log points). The magnitudes of all changes are smaller here. However, the direction of change is identical for each wage variable we examine.

Models 2−3 explore the effects of varying the elasticity of substitution between skilled and unskilled labor. The base model uses the estimated elasticity of 1.441. The Cobb−Douglas case (with an elasticity of substitution equal to 1) shows the largest wage effects. The direction of changes in wages is the same for each case except for the slight decline in overall wage inequality (as measured by the standard deviation of log wages) when the elasticity of substitution is 2.

Models 5 and 6 explore the effects of changes in the utility parameter that determines the intertemporal elasticity of substitution for consumption. The effects of technology change on wages and wage inequality are very similar to the base case model.

Model 7 shows the simulated changes in wages in response to a change in technology when the model is estimated incorrectly as discussed in Appendix B, case 2. In this case, the model is first simulated using the base case technology change specification. Then we estimate parameters of the model using data from the early transition years of the simulated data (under the assumption of constant skill prices and interest rates) and we recalibrate the technology parameters of the model. Using those new parameters, we simulate a constant decline in a_1 producing a 30% decline in low-skill human capital. These results are very similar to the base case

TABLE C-1

Percentage Changes in Wages and Wage Inequality from Initial Steady State to 10 Years in Transition

Simulated economy[a]	Coll.–HS log wage diff.	Mean HS log wage		Mean coll. log wage		Std. deviation of log wages		
		Age 25	Age 50	Age 25	Age 50	HS	College	All
1. Base case	2.02	−7.38	−2.97	4.12	3.05	−2.30	1.47	0.14
2. Elasticity of sub. = 0.5	3.54	−8.61	−3.35	5.00	4.27	−0.33	1.48	1.51
3. Elasticity of sub. = 2	1.33	−6.50	−2.64	3.33	2.64	−2.65	1.66	−0.17
4. Elasticity of sub. = 1	3.99	−9.61	−4.00	6.65	4.15	−1.15	0.63	1.03
5. $\gamma = 0$	2.04	−7.32	−2.99	4.44	3.03	−2.42	1.27	0.03
6. $\gamma = -5$	2.32	−6.74	−3.16	7.68	2.83	−3.54	−0.72	−0.98
7. Misspecified estimation	2.25	−7.29	−3.08	2.65	3.08	−1.70	1.93	0.68
8. σ_ε − 1 std. dev.	0.37	−7.31	−2.31	4.06	2.48	−2.15	1.20	−0.38
9. σ_ε + 1 std. dev.	3.16	−7.38	−3.47	3.89	3.46	−2.53	1.81	0.51
10. Open economy	1.30	−9.76	−2.77	−6.24	3.38	2.51	8.72	4.60
11. 30% decline in HS share over 30 yrs; doubling of low ability share yrs 1–10	4.45	−13.28	−3.53	2.01	3.66	−0.41	2.49	2.05
12. 30% decline in HS share over 30 yrs; doubling of low ability share yrs 1–30	3.30	−12.89	−3.42	−0.44	3.54	−1.15	3.84	1.81
13. 50% decline in HS share over 60 yrs	1.73	−6.24	−2.53	3.81	2.62	−2.04	1.21	0.06
14. 99% decline in HS share over 100 yrs	2.52	−9.69	−3.77	4.67	3.87	−2.99	2.04	0.25

[a] See Notes to Table C-3 for explanations.

TABLE C-2
Percentage Changes in Wages and Wage Inequality from Initial Steady State to 15 Years in Transition

Simulated economy[a]	Coll.–HS log wage diff.[a]	Mean HS log wage		Mean coll. log wage		Std. deviation of log wages		
		Age 25	Age 50	Age 25	Age 50	HS	College	All
1. Base case	2.50	−10.34	−1.60	13.08	1.90	−1.61	0.25	0.29
2. Elasticity of sub. = 0.5	2.23	−5.42	0.50	9.80	1.81	−0.11	0.87	1.05
3. Elasticity of sub. = 2	2.13	−10.51	−1.84	12.97	1.87	−2.01	0.39	0.03
4. Elasticity of sub. = 1	3.22	−7.80	−0.39	9.17	1.33	−0.46	−0.32	0.84
5. $\gamma = 0$	2.53	−10.44	−1.62	13.02	1.89	−1.75	0.04	0.17
6. $\gamma = -5$	2.88	−11.49	−1.87	12.10	1.77	−3.11	−1.94	−0.91
7. Misspecified estimation	2.70	−9.30	−2.14	9.34	2.19	−1.25	1.00	0.85
8. $\sigma_\varepsilon - 1$ std. dev.	0.23	−8.96	−0.96	9.98	1.48	−1.54	0.15	−0.44
9. $\sigma_\varepsilon + 1$ std. dev.	4.20	−11.34	−2.09	15.48	2.24	−1.79	0.44	0.82
10. Open economy	1.61	−7.36	−0.72	14.06	2.54	4.33	7.44	5.15
11. 30% decline in HS share over 30 yrs; doubling of low ability share yrs 1–10	4.94	−15.99	−2.05	15.40	2.44	−0.86	0.70	1.54
12. 30% decline in HS share over 30 yrs; doubling of low ability share yrs 1–30	.451	−16.41	−2.01	13.06	2.30	−0.25	2.46	2.23
13. 50% decline in HS share over 60 yrs	1.74	−8.37	−1.33	10.41	1.59	−1.79	0.40	−0.01
14. 99% decline in HS share over 100 yrs	2.64	−12.95	−1.99	15.30	2.46	−2.44	0.89	0.24

[a] See Notes to Table C-3 for explanations.

TABLE C-3

Percentage Changes in Wages and Wage Inequality from Initial Steady State to 20 Years in Transition

Simulated economy	Coll.–HS log wage diff.	Mean HS log wage		Mean coll. log wage		Std. deviation of log wages		
		Age 25	Age 50	Age 25	Age 50	HS	College	All
1. Base Case	3.81	−10.52	−1.71	15.03	2.56	−0.80	−1.06	0.68
2. Elasticity of sub. = 0.5	2.59	−4.24	1.71	12.15	2.94	0.12	−0.25	0.91
3. Elasticity of sub. = 2	3.62	−11.02	−2.20	14.86	2.49	−1.24	−0.91	0.47
4. Elasticity of sub. = 1	3.77	−7.10	0.21	10.80	1.78	0.25	−1.53	0.93
5. $\gamma = 0$	3.85	−10.65	−1.76	14.95	2.52	−0.97	−1.26	0.55
6. $\gamma = -5$	4.19	−12.01	−2.28	13.97	2.13	−2.57	−3.20	−0.58
7. Misspecified estimation	3.96	−9.43	−2.40	10.87	2.85	−0.77	−0.05	1.24
8. $\sigma_\varepsilon - 1$ std. dev.	1.20	−8.63	−0.82	11.44	1.92	−0.90	−1.04	−0.24
9. $\sigma_\varepsilon + 1$ std. dev.	5.86	−11.92	−2.43	17.84	3.09	−0.88	−0.97	1.37
10. Open economy	2.94	−5.80	−0.02	17.32	3.98	5.37	5.46	5.32
11. 30% decline in HS share over 30 yrs; doubling of low ability share yrs 1–10	5.34	−11.68	−2.09	19.35	3.11	−0.43	−0.88	1.40
12. 30% decline in HS share over 30 yrs; doubling of low ability share yrs 1–30	6.55	−16.95	−2.31	15.76	3.28	0.70	0.81	2.83
13. 50% decline in HS share over 60 yrs	1.92	−7.84	−1.24	10.65	1.92	−1.81	−0.17	−0.09
14. 99% decline in HS share over 100 yrs	3.11	−12.28	−1.79	16.07	3.20	−2.20	0.01	0.22

Notes.

1. The base case is as defined in this appendix. (Note that the rate of decline in a_1 differs from that in the main paper.) Human capital production functions and schooling functions as reported in Tables I and II in the text. $\sigma_1 = 1.441$; $\sigma_2 = 1$. Technology parameter a_1 linearly declines for 30 years yielding a 30% total decline in high school share.

2–4. Elasticity of substitution between high school and college human capital, σ_1, as noted. Other parameters as in the base case.

5, 6. Utility parameter, γ, as noted. Other parameters as in the base case.

7. All parameters for the model were estimated on simulated data as described in Appendix B. The model was simulated with a change in technology, but parameters were estimated under the assumption that technology was stable. The results reported are for a 30 year decline in the newly calibrated a_1 that yields a total decline in high school share of 30%.

8, 9. One standard deviation perturbations of the estimated variance of the nonpecuniary costs of schooling. Other parameters as in the base case.

10. Open economy: fixed after-tax interest rate at 5%.

11, 12. Same as base case with a doubling in the cohort size for the first 10 or 30 years of the transition period.

13, 14. Same as base case with variations in the length of technology change and the total change in high school share.

results, suggesting that misspecifying our estimation procedure in this way (failing to account for changing skill prices in estimation) does not lead to incorrect conclusions.

Models 8 and 9 explore the effects of changing the variance in nonpecuniary costs to schooling. Larger variances lead to slightly larger increases in the college−high school wage differential, since the supply of skill is more inelastic. The qualitative changes in wages and inequality are unchanged.

In model 10, we show the effects of an open capital market where the interest rate is held constant at its initial steady state level. In Table C-1, the 10 year effects on the standard deviation of log wages are all positive, suggesting a rise in within group inequality as well as overall inequality. The decline in wages for young college graduates observed in Table C-1, reverses itself after 15 years as seen in Table C-2. Thus, the only major difference between the open-economy case and the base case is the change in the standard deviation of log wages for high school graduates.

Models 11 and 12 show the effects of increasing the proportion of workers from the lowest quartile of ability. This substantially depresses the wages of young high school workers and slightly depresses the mean wages for young college workers. The college−high school wage gap and the standard deviation of log wages rise more over each of the periods we examine. Thus, huge increases in the proportion of low ability workers (either from migration or other cohort differences) cause inequality to increase more, though the effects are rather small compared to the magnitude of the change in cohort composition.

Models 13 and 14 show the effects of extending the duration of technology change. Qualitatively, the results are the same. The college−high school wage gap rises as does overall wage inequality. The gap also rises more for young workers than old workers as in the base case.

All of the variations in our model we consider produce the same qualitative conclusions presented in the text of the paper. With the onset of skill-biased technological change, the college−high school wage gap rises. Overall wage inequality (as measured by the standard deviation of log wages) rises in most specifications. With only a few exceptions, within-group wage inequality rises for college graduates but falls for high school graduates. Wage profiles typically steepen for high school graduates and flatten for college graduates.

REFERENCES

1. A. Auerbach and L. Kotlikoff, "Dynamic Fiscal Policy," Cambridge University Press, Cambridge, UK, 1987.
2. D. Autor, L. Katz, and A. Krueger, "Computing Inequality: Have Computers Changed the Labor Market?," NBER Working Paper 5956, 1997.

3. P. Beaudry and H. Green, ''Cohort Patterns in Canadian Earnings: Assessing the Role of Skill Premia Inequality Trends,'' NBER Working Paper 6132, 1997.

4. Y. Ben Porath, The production of human capital and the life cycle of earnings, *J. Polit. Econ*. **75** (1967), pt. 1, 352–365.

5. J. Bound and G. Johnson, Changes in the structure of wages in the 1980s: An evaluation of alternative explanations, *Amer. Econ. Rev*. **82** (1989), 371–392.

6. C. Brown, A model of optimal human-capital accumulation and the wages of young high school graduates, *J. Polit. Econ*. **84** (1976), 299–316.

7. M. Browning, L. Hansen, J. Heckman, and C. Taber, Calibrating and estimating general equilibrium models, *in* ''Handbook of Macroeconomics'' (J. Taylor and M. Woodford, Eds.), North-Holland, Amsterdam, 1998.

8. S. Cameron and J. Heckman, The dynamics of educational attainment for blacks, whites and hispanics, unpublished manuscript, 1997.

9. S. Cameron and J. Heckman, Lifecycle schooling and educational selectivity: Models and choice, *J. Polit. Econ*. April, (1998), 262–333 (first presented at the University of Wisconsin, June 1990).

10. F. Caselli, ''Technological Revolutions,'' Univ. of Chicago Business School, June 1997.

11. C. Clotfelter, R. Ehrenberg, M. Getz, and J. Siegfried, ''Economic Challenges in Higher Education,'' Univ. of Chicago Press, Chicago, 1991.

12. J. Davies and J. Whalley, Taxes and capital formation: How important is human capital? *in* ''National Saving and Economic Performance'' (B. Bernheim and J. Shoven, Eds.), Univ. of Chicago Press, Chicago, 1991.

13. R. Freeman, ''The Overeducated American,'' Basic Books, New York, 1976.

14. J. Greenwood and M. Yorukoglu, ''1974,'' Carnegie–Rochester Series on Public Policy, 1997

15. D. Hamermesh, ''Labor Demand,'' Princeton Univ. Press, Princeton, 1993.

16. J. Heckman, Estimates of a human capital product on function embedded in a life-cycle model of labor supply, *in* ''Household Production and Consumption'' (N. Terleckyj, Ed.), pp. 227–264, Columbia Univ. Press, New York, 1975.

17. J. Heckman, A life-cycle model of earnings, learning, and consumption, *J. Polit. Econ*. **84** (1976), pt. 2, S11–S44.

18. J. Heckman, What has been learned about labor supply in the past twenty years? *Amer. Econ. Rev. Papers Proc*. **83** (1993), 116–121.

19. J. Heckman, ''Should the U.S. Have a Human Capital Policy and If So, What Should It Be?,'' Gilbert Lecture, Univ. of Rochester, April 1996.

20. J. Heckman and P. Klenow, Is there underinvestment in human capital?, unpublished manuscript, Univ. of Chicago, 1997.

21. J. Heckman, L. Lochner, and C. Taber, Formulating and estimating dynamic general equilibrium models to evaluate policies that promote skill formation, Marschak Lecture, Econometric Society Meetings, Hong Kong, July 25, 1997.

22. J. Heckman and G. Sedlacek, Heterogeneity, aggregation and market wage functions: An empirical model of self-selection in the labor market, *J. Polit. Econ*. **98** (1985), 1077–1125.

23. J. Heckman and P. Todd, 40 years of Mincer earnings functions, unpublished manuscript, Univ. of Chicago, 1997.

24. G. Johnson, The demand for labor by educational category, *Southern Econ. J*. **37** (1970), 190–204.

25. T. Kane, College entry by blacks since 1970: the role of college costs, family background and the returns to education, *J. Polit. Econ*. **102** (1994), 878–911.

26. L. Katz and K. Murphy, Changes in relative wages, 1963–1987: Supply and demand factors, *Quart. J. Econ*. **107** (1992), 35–78.

27. M. Keane and K. Wolpin, The career decisions of young men, *J. Polit. Econ*. **105** (1997), 473–522.

28. P. Krussell, L. Ohanian, J. V. Rios-Rull, and G. Violante, Capital skill complementarity and inequality: A macroeconomic analysis, unpublished manuscript, Univ. of Pennsylvania, 1996.

29. F. Kydland, Business cycles and aggregates labor market fluctuations, *in* ''Frontiers in Business Cycle Research'' (T. Cooley, Ed.), Princeton Univ. Press, Princeton, NJ, 1995.

30. P. Lipton, J. Sacks, and J. Poterba, Multiple shooting in rational expectations models, *Econometrica* **50** (1982), 1329–1334.

31. R. E. Lucas, On the mechanics of economic development, *J. Monetary Econ.* **22** (1982), 3–42.

32. T. MaCurdy and T. Mroz, Measuring microeconomic shifts in wages from cohort specification, unpublished manuscript, Stanford Univ., 1995.

33. C. Meghir and E. Whitehouse, The evolution of wages in the United Kingdom: Evidence from micro data, *J. Labor Econ.* **14** (1996), 1–25.

34. J. Mincer, On the job training: Costs, returns and some implications, *J. Polit. Econ.* **70** (1962), Suppl. 5, pt. 2, 50–79.

35. J. Mincer, ''Schooling, Experience and Earnings,'' Columbia Univ. Press, New York, 1974.

36. S. Ortigueira and M. Santos, ''On Convergence in Endogenous Growth Models,'' Discussion Paper 9409, ITAM, 1994.

37. J. Pechman, ''Federal Tax Policy,'' 5th ed., Brookings Institution, Washington, DC, 1987.

38. J. Poterba, The rate of return to corporate capital and factor shares: New estimates using data revised national income and capital stock, unpublished manuscript, MIT, 1997.

39. S. Rosen, A theory of life earnings, *J. Polit. Econ.* **84** (1976), pt. 2, S45–S57.

40. A. D. Roy, Some thoughts on the distribution of earnings, *Oxford Econ. Papers* **3** (1951), 135–146.

41. M. Sattinger, Assignment models of the distribution of earnings, *J. Econ. Lit.* **31** (1993), 831–880.

42. A. Siow, Occupational choice under uncertainty, *Econometrica* **52** (1984), 631–645.

43. H. Uzawa, Optimum technical change in an aggregative model of economic growth, *Int. Econ. Rev.* **6** (1965), 18–31.

44. G. Violante, Equipment investment and skill dynamics: A solution to the wage dispersion puzzle, unpublished manuscript, U.C. London, 1996.

45. R. Willis and S. Rosen, Education and self-selection, *J. Polit. Econ.* **87** (1979), pt. 2, S7–S36.

46. A. Bartel and N. Sicherman, ''Technological Change and the Skill Acquisition of Young Workers,'' National Bureau of Economic Research, Working Paper S107, 1995.