Review of Economic Dynamics 12 (2009) 168–182

Contents lists available at ScienceDirect

Review of Economic Dynamics

www.elsevier.com/locate/red

Education, sectoral composition, and growth

Joseph P. Kaboski *,[1]

Ohio State University, USA

ARTICLE INFO

Article history:
Received 23 June 2006
Revised 29 July 2008
Available online 5 August 2008

JEL classification:
O2
O3
I2

Keywords:
Economic growth
Human capital
Structural transformation
Education

ABSTRACT

Growth accounting exercises using standard human capital measures are limited in their ability to attribute causal effects and to explain growth. This paper develops a model of growth and schooling consistent with these decompositions but with less unexplained growth. The theory distinguishes between three different sources of education gains: (1) supply shifts, (2) skill-biased technical change increasing demand within industries/occupations, and (3) skill-biased technical change caused by the introduction of new skill-intensive industries/occupations. The third source leads to the large sectoral shifts and the largest growth effects. Quantitatively, schooling contributions account for 24 percent of wage growth, with both the direct (i.e., supply driven) causal contribution of schooling and the indirect causal (i.e., technology induced) contribution playing substantial roles.

1. Introduction

Is schooling an important and direct driving force of economic growth? One viewpoint in the literature, including theory and growth accounting decompositions, attributes large direct gains in output to schooling and human capital investment (e.g. Denison, 1974; Jorgenson et al., 1987; Krueger and Lindahl, 2001; Lucas, 1988; Young, 1995). More recently, another viewpoint (e.g. Bils and Klenow, 2000; Klenow and Rodriguez-Clare, 1997) has claimed that the causal connection from schooling investment to growth explains very little of overall growth.[2] This proponents of this view argue persuasively that the "contributions" from growth accounting decompositions cannot be interpreted as causal, and so models are necessary to distinguish causal forces. This paper develops a model of human capital that does precisely this. The model distinguishes between causal forces in the growth of education and yet is also consistent with salient features of growth accounting exercises on U.S. wages from 1940 to 1970.

Section 2 highlights the limitations of growth accounting exercises and standard human capital measures—one, a measure of human capital based on Mincerian returns and average schooling and, the other, a more flexible schooling-based measure. These decompositions though not causal, are nonetheless informative. In short, decompositions show that reallocations of workers toward higher levels of schooling are not positively associated with growth. However, allocations toward higher paying and higher skill occupations and industries are strongly associated with growth. In the standard models, sector plays no role.

* Address for correspondence: Ohio State University, 415 Arps Hall, Columbus, OH, USA.
E-mail address: kaboski.1@osu.edu.

[1] I am grateful for helpful comments received from George Alessandria, Francisco Buera, Bill Dupor, Paul Evans, Kathryn Moran, Nancy Stokey, Bruce Weinberg, the editor, and two anonymous referees.

[2] Another similar debate exists about the relative importance of human capital and TFP in explaining cross-country variation in output per capita at a point in time (e.g. Hall and Jones, 1999; Hendricks, 2002; Kaboski, 2007; Manuelli and Seshadri, 2005; Mankiw et al., 1992).

1094-2025/$ – see front matter
doi:10.1016/j.red.2008.07.003

In Section 3, I present a more nuanced concept of human capital, where workers with heterogeneous skills sort themselves into different sectors (empirically proxied by industry or occupation). The returns to skill endogenously respond to the movements in supply and demand, and the model can therefore distinguish between different causal forces behind schooling increases. These different causal forces affect not only the causal contribution of schooling, but also whether or not rising schooling levels coincide with periods of high wage growth. Thus, the model can explain the relationship between increases in schooling and contemporaneous growth.

In the model, there are three forces that can lead to changes in schooling levels. They are distinguished by their effects on output, the returns to schooling, and the amount of schooling that can be explained by reallocations of labor across sectors:

- An outward supply shift is modeled as a fall in the cost of schooling. These supply shifts cause relatively small output gains, declining returns to schooling, and a reallocation of labor toward low-skill sectors.
- The second force is a type of skill-biased technical change (SBTC). This SBTC introduces new industries and occupations with high demand for skills, and therefore involves structural transformation across sectors. This SBTC is consistent with empirical findings that newer, higher technology industries, occupations and products are more skill-intensive.[3] These new skill-intensive sectors leads to strong growth, rising returns to schooling, and important reallocations of workers toward high-skill, high-wage sectors.
- The third force is second type of SBTC that is simply a change in the relative productivity of skilled workers within existing sectors. It could be motivated by a fall in the relative price of physical capital for skilled workers across all sectors in a story similar to Krusell et al. (2000). It is consistent with evidence that recent SBTC is not explained by sectoral shifts, but by a pervasive increase in the demand for skill within industries and occupations.[4] I assume technologies that increase the productivity of the highly-skilled in these sectors, also decrease the productivity of the less skilled, and so have relatively little net productivity effects. Hence, this second type of skill-biased technical change leads to small output gains, increasing returns to schooling, and rising schooling levels that is explained almost exclusively by growth in schooling within sectors, rather than reallocations toward high-skill sectors.

The focus on skill-biased technical change contributes to a large literature on the sources and growth effects of SBTC in two novel ways.[5] It presents a more plausible argument of capital-skill complementarity, in which certain capital goods are complementary with skilled labor, those with a falling relative price. More importantly, it links skill-biased technical change to observable sectoral shifts. Thus, it can be viewed as adding a human capital component to the literature on structural transformation.[6]

The quantitative analysis of Section 4, calibrates the three forces above to match changes in average years of schooling, changes in Mincerian returns to schooling, and the amount of schooling growth explained by sectoral reallocations in each decade. Residual wage growth is substantially smaller and also less volatile than in the typical human capital models. The calibration identifies that pre-1970 growth was characterized by rapid introduction of high skill sectors, technical change within sectors that is biased toward low skill workers, and an outward shift in the supply of education, while the post-1970 has been characterized by slower introduction of new high skill sectors, SBTC within sectors, and a backward shift in the supply schedule of education (particularly in the 1980s and 1990s).

Within such an interpretation, on average 24 percent of wage growth is attributed to increases in schooling, although the sources of this vary by time period. The model attributes 33 percent of wage growth in the earlier period and −13 percent in the later period to the direct (i.e. supply shift) contribution of education. The indirect contributions that come from educational increases ultimately driven by technology are −21 and 52 percent, respectively. These indirect contributions are policy relevant to the extent that frictions or policies might prevent workers from endogenously responding to changes in technology.

2. Limitations of existing models and methods

This section examines the limitations of existing models and methods. The models discussed are two production functions based on schooling that augments the efficiency units of human capital that are representative of the literature. The methods involve measuring the growth of measured human capital stocks and attributes it to exogenous increases in education investments. I first introduce the data, then the two production functions, and finally the growth analyses. Finally, alternative growth decompositions are used to motivate the sectoral model.

[3] See Goldin and Katz (1998), Machin and Van Reenen (1998), Murphy and Welch (1993), and Xiang (2005).

[4] See Berman et al. (1994, 1998), Juhn et al. (1993), Katz and Murphy (1992), and Machin and Van Reenen (1998).

[5] Important contributions to the macrogrowth literature include Acemoglu (1998), Caselli (1999), Galor and Moav (2000), Galor and Tsiddon (1997), Galor and Weil (2000), Greenwood and Yorukoglu (1997), and Krusell et al. (2000).

[6] Recent contributions include Echevarria (1997), Acemoglu and Guerrieri (2005), Gollin et al. (2002), Kongsamut et al. (2001), Matsuyama (1992, 2002), and Ngai and Pissarides (2004).

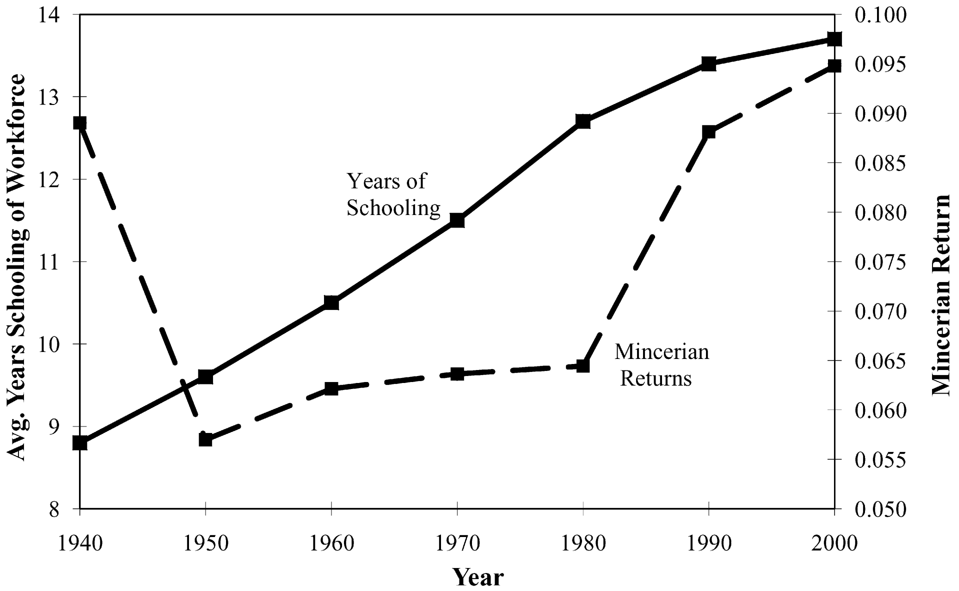

Fig. 1. Average schooling levels and Mincerian returns, 1940–2000.

2.1. Data

The analyses are based on integrated use public microdata series (IPUMS) census data of individuals that contains annual hours, annual labor earnings, industry, occupation, years of schooling, age and sex. Schooling and earnings are first available in the 1940 U.S. census, so I focus on the seven decennial censuses from 1940 to 2000. Labor earnings exclude benefits, and wages are deflated hourly wages calculated as *earnings/hours*. Average years of schooling are hours-weighted averages. The Mincerian return is the estimated coefficient on schooling from a regression of log wages on years of schooling, years of experience, and squared years of experience à la Mincer (1974). As is standard, years of experience is constructed as *age* − *years of schooling* − 6, and the Mincerian regressions include only full-time males in order to abstract from endogenous labor supply issues.

For the decompositions, I aggregate schooling, industry, and occupation into broader categories, since classification errors are particularly high for industry and occupation at disaggregate levels.[7] The standard nine broad occupational categories are used for occupation,[8] while for industry, I grouped the standard sixteen broad industries into nine industry groups in order to maintain comparability.[9] The schooling classification also involves standard nine schooling level groupings.[10] Appendix A gives further details on the data.

Fig. 1 shows average schooling levels (solid line) and returns to schooling (dashed line) in the U.S. workforce from 1940 to 2000. Average educational levels rise throughout the period, and the growth is fairly stable over time. On average, schooling levels rose 0.90 years per decade during the high economic growth from 1940 to 1970, and 0.73 during the slower post-1970 growth period. In contrast to schooling levels, the returns to schooling fluctuate over time. The measured return is the estimated coefficient on schooling from a regression of log wages on years of schooling, years of experience, and squared years of experience à la Mincer (1974). Both the levels and time path of returns replicate the well-known pattern: a large drop from 1940 to 1950, then a slow increase followed by a faster increase after 1980.

2.2. Standard models

Consider the following production function, where output is Cobb–Douglas in physical $K(t)$ and human capital $H(t)$:

$$Y(t) = A(t)K(t)^{\alpha}\left[H(t)\right]^{1-\alpha}.$$

[7] My results presented are robust to using the more detailed classifications, however.

[8] These are: (i) laborers, (ii) farmowners & farmworkers, (iii) service workers, (iv) operatives, (v) craftsmen & foremen, (vi) salesworkers, (vii) clerical workers, (viii) managers, officials, & proprietors, and (ix) professional & technical.

[9] The nine groups were grouped to keep industries with similar average educational levels together, but create groups of comparable size. The groupings are: (i) agriculture/forestry/fishing, mining and personal services (grouped); (ii) nondurable manufacturing; (iii) construction; (iv) transportation, utilities, and entertainment and recreational services (grouped); (v) durable manufacturing; (vi) wholesale trade; (vii) retail trade; (viii) telecommunications, business/repair services, public administration and finance/insurance/real estate (grouped); and (ix) professional services. The results are both qualitatively and quantitatively similar regardless of whether this nine-industry grouping or the original 16-industry classification is used.

[10] The upper bounds on years of schooling for the nine standard categories are: (i) 0 years, (ii) up to 4 years, (iii) up to 8 years, (iv) 9 years, (v) 10 years, (vi) 11 years, (vii) 12 years, (viii) less than 16 years, and (ix) 16 years or more.

Human capital can be broken out into the amount of labor $L(t)$ and the average human capital per unit of labor $h(t)$, $H(t) = L(t)h(t)$. In perfect competition, marginal products of labor and capital equal their factor prices, $w(t)$ and R, so that:

$$w(t) = (1-\alpha)A(t)\left[\frac{K(t)}{h(t)L(t)}\right]^{\alpha} h(t),$$
$$R = \alpha A(t)\left[\frac{K(t)}{h(t)L(t)}\right]^{\alpha-1}.$$

Assuming further that capital accumulation keeps the rental rate of capital constant, yields the following wage equation:

$$\begin{aligned} w(t) &= (1-\alpha)A(t)^{\frac{1}{1-\alpha}}\left[\frac{R}{\alpha}\right]^{\frac{\alpha}{\alpha-1}} h(t) \\ &= \omega h(t), \end{aligned} \tag{1}$$

where I have defined $\omega \equiv (1-\alpha)A(t)^{\frac{1}{1-\alpha}}[\frac{R}{\alpha}]^{\frac{\alpha}{\alpha-1}}$ as the price on efficiency units of human capital.

2.2.1. Mincerian model

One typical approach is to construct human capital stocks using the Mincerian relationship and average years of schooling.[11] A production function in which each (average) year of schooling $s(t)$ yields an $r(t)$ percentage increase in human capital is often motivated by the relationship:

$$h(t) = h_0 e^{r(t)s(t)}.$$

Inserting this into Eq. (1) yields the following wage equation:

$$w(t) = (1-\alpha)A(t)^{\frac{1}{1-\alpha}}\left[\frac{R}{\alpha}\right]^{\frac{\alpha}{\alpha-1}} h_0 e^{r(t)s(t)}.$$

Taking logs and differentiating yields a growth accounting equation:

$$\frac{dw/dt}{w} = r\frac{ds}{dt} + s\frac{dr}{dt} + \frac{1}{1-\alpha}\frac{dA}{A},$$

to which a discrete approximation is:

$$\frac{\Delta w_t}{\overline{w}_t} = \bar{r}_t \Delta s_t + \bar{s}_t \Delta r_t + \frac{1}{1-\alpha}\frac{\Delta A_t}{\overline{A}_t}. \tag{2}$$

Here variables with the upper bar indicate the time average over t and $t+1$, while Δ indicates the difference between $t+1$ and t. The first two terms capture the wage growth explained by increases in human capital. In this setup, capital accumulation is endogenous, and the endogenous capital accumulation is attributed to the changes in human capital or technical change.

As can be seen from Fig. 1, although growth in s_t is relatively stable across decades, the Mincerian return r_t is volatile. Using the data in Fig. 1, one can easily construct the first two terms of Eq. (2). For the sake of brevity, I do not present the data in table format, but merely discuss the results. The growth explained by these first two terms is large on average (exceeding actual wage growth, 20.5 to 14.6%), but is unfortunately negatively correlated (−0.91) with actual wage growth. Given this negative correlation, the volatility of *residual* wage growth is greater than that of actual wage growth with standard deviations of 31.4 and 14.6%, respectively. That is, although the decomposition more than explains long-run wage growth, it has apparently nothing to say about decade-by-decade variation in wages, which is a problem.

This model illustrates the volatility of the Mincerian return and the problem of the low correlation between schooling investment and actual growth, but restricts the analysis in several ways. First, h_0, the human capital of a person with no education, is fixed so that changes in $r(t)$ imply large changes in absolute levels of efficiency units for workers with schooling, rather than just changes in relative human capital. Second, it does not allow for differences in human capital across agents. Finally, it insists on a linear wage vs. schooling relationship.

2.2.2. More general model

A more general model would allow relative wages to vary non-parametrically across schooling levels and also over time. I address this by modeling differentiated (imperfectly substitutable) human capitals indexed by $n = 1 \ldots N$ different schooling categories:

$$h(t) = F\big[l(s_1,t),\ldots,l(s_N,t)\big].$$

[11] Bils and Klenow (2000) is an important example.

Here, $F(\cdot)$ is homogeneous of degree one, $h(t)$ is again the average human capital per worker, $l(s_n,t)$ is the fraction of workers with education s_n at time t. Their wage $w(s_n,t)$ is again their marginal product:

$$w(s_n,t)=\omega\frac{\partial F}{\partial l(s_n,t)},$$

but the average wage is again:

$$\begin{aligned} w(t)&=\omega F\bigl[l(s_1,t),\dots,l(s_N,t)\bigr]\\ &=\omega\sum_n\frac{\partial F}{\partial l(s_n,t)}l(s_n,t),\end{aligned}$$

where the second line follows from the homogeneity of $F(\cdot)$.

Substituting in $\omega=(1-\alpha)A(t)^{\frac{1}{1-\alpha}}[\frac{R}{\alpha}]^{\frac{\alpha}{\alpha-1}}$, taking logs, and again totally differentiating with respect to time yields the growth accounting equation:

$$\frac{dw(t)/dt}{w(t)}=\frac{\sum_n\frac{\partial F}{\partial l(s_n,t)}\frac{dl(s_n,t)}{dt}}{h(t)}+\frac{1}{1-\alpha}\frac{dA(t)/dt}{A(t)}.$$

Since, $\frac{\partial F}{\partial n}$ is held fixed at a point in time, fluctuations in $r(t)$ (or any other measure of relative wages) over time do not imply changes in wages, but only changes in the marginal value of shifting people across levels of schooling, s_n. After multiplying both the numerator and denominator of the first term on the right-hand side by the wage per efficiency unit ω, one can rewrite the accounting equation in terms of wages. In discrete form:

$$\frac{\Delta w_t}{\overline{w}_t}=\frac{\sum_n\overline{w_t}(s_n)\Delta l_t(s_n)}{\overline{w}_t}+\frac{1}{1-\alpha}\frac{\Delta A_t}{\overline{A}_t},\tag{3}$$

where $\overline{w_t}(s_i)$ indicates the wage of workers with schooling s_n, averaged across t and $t+1$.[12]

2.3. Decompositions

This section interprets Eq. (3) as a decomposition. I then use this result to introduce other decompositions with respect to industry and occupation. While these decompositions are *ad hoc* at this point, they will be more successful in explaining decade-to-decade changes in the data, and so they are of interest. Ultimately, these decompositions motivate a model developed in Section 3 that is explicit about sector.

To see that Eq. (3) is a decomposition, note that algebraically, the change in wages can be expressed:

$$\frac{\Delta w_t}{\overline{w}_t}=\frac{\sum_n\overline{w_t}(s_n)\Delta l_t(s_n)}{\overline{w}_t}+\frac{\sum_n\Delta w_t(s_n)\bar{l_t}(s_n)}{\overline{w}_t}.\tag{4}$$

The first right-hand side term is the change in wages predicted from the compositional change of the workforce, and is identical to the first term in Eq. (3). This is the "between" component of the decomposition—wage growth explained by movement of workers between schooling levels. Similarly, the second right-hand side terms in the two equations must also be equal. These residuals in the decomposition are the "within" component of the growth decomposition—growth in wages within given levels of schooling.

The fact that growth accounting equations are simple decompositions highlights two known limitations of growth accounting methods. First, as merely decompositions, they cannot identify true sources of growth, or causally distinguish schooling-driven growth from technology-driven growth (see Klenow and Rodriguez-Clare, 1997). Second, although these decompositions identify total factor productivity (TFP) innovations within the context of these models, they identify them as a residual. These TFP innovations are unexplained measures of our ignorance as economists.

Decompositions are nonetheless interesting moments in the data, and one can perform a parallel decomposition with respect to industry (or occupation) rather than schooling categories. The models above have no concept of sector, and so these decompositions are *ad hoc* at this point, but the model in the next section justifies such a decomposition. Denoting sector as i_j (for industry $j=1\dots J$), we have:

$$\frac{\Delta w_t}{\overline{w}_t}=\frac{\sum_j\overline{w_t}(i_j)\Delta l_t(i_j)}{\overline{w}_t}+\frac{\sum_j\Delta w_t(i_j)\bar{l_t}(i_j)}{\overline{w}_t}.\tag{5}$$

[12] The standard growth accounting representation of Eq. (3) is:

$$\frac{\Delta w_t}{\overline{w}_t}=\sum_i\overline{\sigma}(s_i)\frac{\Delta l_t(s_i)}{l_t(s_i)}+\frac{1}{1-\alpha}\frac{\Delta A_t}{\overline{A}_t},$$

where $\sigma(s_i)$ is the share of wages going to workers of type s_i. However, Eq. (3) leads more easily into (4), which shows the "between" and "within" interpretation of the decomposition.

Table 1
Growth decompositions of U.S. data, 1940–2000

Time period:	Explaining wage growth (per decade)							Explaining schooling growth (per decade)				
	Avg. percent wage growth	Schooling decomposition		Industry decomposition		Occupation decomposition		Avg. years of schooling growth	Industry decomposition		Occupation decomposition	
		"Explained" growth[a]	Residual	"Explained" growth[a]	Residual	"Explained" growth[a]	Residual		"Explained" growth[a]	Residual	"Explained" growth[a]	% of total
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)	(x)	(xi)	(xii)
Pre-1970	34%	4%	30%	7%	28%	7%	27%	0.90	0.25	0.65	0.32	0.58
Post-1970	10%	5%	5%	0%	9%	2%	8%	0.73	0.09	0.64	0.14	0.59
1940–2000	21%	5%	17%	3%	18%	4%	21%	0.82	0.17	0.65	0.24	0.58
Standard deviation[b]	14.6%	0.1%	15.2%	4.5%	11.4%	3.9%	12.0%	0.31	0.10	0.26	0.11	0.24
Time correlation w/actual wage growth[b]		−0.48	1.00	0.78	0.97	0.74	0.98		0.85	0.26	0.73	0.27

[a] "Explained" growth is the growth in wages (or schooling) explained by changes in quantities of labor inputs at given wages (or given schooling levels, respectively). See Eqs. (4)–(6).

[b] The standard deviation is across the values of the six decade-to-decade values, while the correlation is the correlation of the values of actual wage growth and explained wage (or schooling) growth across these six decade-to-decade values.

Finally, the increase in average schooling can be similarly decomposed into the increase explained by changes in the sectoral composition of the workforce, and increases in schooling levels with sectors:

$$\Delta s_t = \sum\nolimits_j \overline{s_t}(i_j)\Delta l_t(i_j) + \sum\nolimits_j \Delta s_t(i_j)\overline{l_t}(i_j). \tag{6}$$

Although, one can construct all of the summations in (4)–(6) directly, I follow growth accounting terminology and refer to the first term (the "between" component) as the "explained growth" and the second term (the "within" component) as the "residual."

2.4. Empirical results

Table 1 presents the empirical results for U.S. wages and schooling, 1940–2000. These results highlight the difficulties with the model presented in 2.2.2, the limitations of decomposition (4), and the insights of decompositions (5) and (6). I discuss each column in turn.[13]

- Examining column (i), wage growth averages 21 percent per decade over the six decades but shows considerable variation over time with a standard deviation of 14.6 percent. In particular, during the high growth of 1940–1970, wage growth is over three times higher on average than the during low growth period of 1970–2000.[14]
- Columns (ii) and (iii) show the results for the decomposition in (4). The wage growth explained by schooling, while high, shows very little variation (standard deviation) and is again negatively correlated (−0.48) with actual wage growth. Indeed, in the 1970s (not shown) the wage growth "explained" by the schooling decomposition exceeds actual wage growth. Moreover, given the negative correlation, the standard deviation of residual/TFP growth (15.2 percent) again exceeds that of actual wage growth as in the Mincerian model.
- Columns (iv) through (vii), based on (5), show the results of the wage decompositions with respect to industry or occupation. Explained growth using these measures of sector are closely associated with observed wage growth. In column (iv), the between-industry growth is highly correlated with actual growth (0.78), averaging 7 percent per decade before 1970, and 0 percent after 1970. The relationship for occupation in column (vi) is nearly as strong with a correlation of 0.74 between explained and actual growth and a drop from 7 percent to 2 percent from the pre- to post-1970 period.
- Columns (viii)–(xii) present the schooling decompositions based on (5). Industry and occupation reallocations are able to explain a substantial fraction of the growth in schooling before 1970 (0.25 and 0.32 years, respectively, out of 0.90 years), but little after 1970 (0.09 and 0.14 out of 0.73 years). Moreover, the amount of schooling growth explained by industry or occupation decompositions is quite highly correlated with wage growth (0.85 and 0.73, respectively).

The models in Sections 2.2.1 and 2.2.2 have no concept of sector, and offer no way of interpreting sectoral decompositions, except perhaps as approximations to the schooling decompositions that yield quite different results. The results show

[13] These results are quite robust to alternative decompositions including creating industry-occupation classes or using three-digit industry or occupation classes. Controlling for experience by subdividing groups by experience (e.g. schooling-experience classes) also produced similar results except that explained growth from schooling was higher (6 vs. 5 percent) and less inversely correlated with actual wage growth (−0.21 vs. −0.48).

[14] Although wage growth was much greater in the 1990s (averaging 1.5 percent/year) than in the 1970s (0.1 percent) or 1980s (1.2 percent), it was still much lower than wage growth from 1940–1970, which averaged 3.0 percent/year. The same is true of output/worker.

that decompositions based on industry and/or occupations appear to be more informative about the sources and time variation of growth compared to decompositions solely based on schooling levels. Therefore, the next section introduces a model that is consistent with these decompositions and with sectoral allocations playing a key role in growth.

3. Model

In the model below, production involves differentiated sectors and differentiated levels of skill that are endogenously chosen through the choice of schooling. The level and distribution of skills will be affected by three factors: the direct cost of schooling that governs the supply of skill and the two skill-biased technology factors that determine the demand for skill. Mincerian returns and relative wages fluctuate endogenously, and schooling increases can accompany either slow or rapid growth.

3.1. Firms

Final output is produced from aggregating the output $\tilde{x}$ of a continuum of sectors (indexed by their level of complexity i).[15] These goods are produced by a continuum of heterogeneous workers using heterogeneous capital. Capital is specific to the type of worker (so both workers and capital are indexed by their skill level h):

$$Y = \tilde{A}\left(\int_{\underline{I}}^{\bar{I}} \tilde{x}(i)^{1-\tilde{\mu}}\, di\right)^{\frac{1}{1-\tilde{\mu}}},$$

$$\tilde{x}(i) = \int_0^\infty \tilde{a}(i,h)k(i,h)^\alpha l(i,h)^{1-\alpha}\, dh. \tag{7}$$

Here $l(i,h)$ indicates the amount of labor with human capital (or skill) level h that works in sector i. By normalizing the total workforce to unity, $l(i,h)$ becomes a density function. Likewise, $k(i,h)$ is the amount of capital specific to h workers in sector i. The level of productivity depends on three factors: a sector- and skill-neutral productivity parameter $\tilde{A}$, a sector- and skill-specific productivity function $\tilde{a}(i,h)$, and I, the maximum complexity level of existing sectors. Notice that the output produced by workers of different skills is perfectly substitutable only if they work in the identical sector.

If capital is fungible across (i,h) types, then production can be simplified (see Appendix B) and expressed as a function of aggregate capital K and aggregate labor services L:

$$Y = \tilde{A}K^\alpha L^{1-\alpha}, \tag{8}$$

$$L \equiv \left(\int_{\underline{I}}^{I} x(i)^{1-\mu}\, di\right)^{\frac{1}{1-\mu}},$$

$$x(i) = \int_0^\infty a(i,h)l(i,h)\, dh. \tag{9}$$

The transformed positive function $a(i,h)$ represents the productivity that a worker of human capital level h has in a sector of complexity level i. The transformed productivity functions now incorporate both inherent productivity differences and differences in capital per worker, which may be caused by the price of complementary capital goods differing across sectors and/or workers. I make the following three assumptions:

$$\text{Assumption 1:}\quad \frac{\partial a(i,h)}{\partial h} > 0,$$

$$\text{Assumption 2:}\quad \frac{\partial^2 \log a(i,h)}{\partial i \partial h} > 0,$$

$$\text{Assumption 3:}\quad \frac{\partial^2 \log a(i,h)}{\partial h^2} < 0.$$

Assumption 1 is that workers with higher levels of skills have an *absolute advantage* over less skilled workers, and ensures that wages will be increasing in skill level. Assumption 2 is a statement of *comparative advantage*; workers with higher levels of skill have a comparative advantage in more complex sectors. Assumption 2 is needed to ensure a sorting equilibrium, in which more skilled laborers work in more complex sectors. This sorting is consistent with the fact that average levels of

[15] In the context of the model, sectors are merely elements of the production process that are imperfectly substitutable and use different production technologies. Industries and occupations are reasonable empirical proxies for these technologies.

schooling differ substantially across industries and occupations (see Juhn et al., 1993). Assumption 3 is an assumption of *log diminishing returns*—in any given sector, incremental gains in skill yield increasingly smaller percentage gains in productivity. This assumption allows the returns to skill to decrease as the skill levels of the workforce increase.

A representative competitive firm takes the wage profile and the rental rate as given and maximizes profits. Since I am interested primarily in the labor market, I further simplify by assuming a fixed world rental price of capital R. Given the optimum choice of capital, production can be expressed:

$$Y = A(R)L, \tag{10}$$

$$A(R) \equiv \tilde{A}^{\frac{1}{(1-\alpha)}} \left(\frac{R}{\alpha}\right)^{\frac{\alpha}{(\alpha-1)}}. \tag{11}$$

The first-order condition for labor for a firm taking the rental price and wage profile as given:

$$w(i,h) = (1-\alpha)A(R)\left[\frac{x(i)}{L}\right]^{-\mu} a(i,h). \tag{12}$$

There is a component of wages, $(1-\alpha)A(R)$, that is common to all workers and similar to the ω defined in (12), but the worker-specific component, $[\frac{x(i)}{L}]^{-\mu}a(i,h)$, depends on both i and h and not simply a number of efficiency units as in the standard model. As the output of a specific sector relative to total output, $\frac{x(i)}{L}$, decreases, the wage increases.

3.2. Agents

I assume a continuum of agents who are heterogeneous across a single dimension, θ, which represents inherent ability. The parameter θ is distributed with positive density across a bounded interval $[\underline{\theta}, \bar{\theta}]$ according to a continuously differentiable density function $g(\theta)$ and is constant over time. Together inherent ability θ and schooling s produce skill h.

Given their inherent ability θ and the direct cost of schooling η which is proportionate to the indirect cost of schooling (i.e. $w(i,h)$), agents simply choose their schooling level s and their sector i to maximize lifetime income:

$$\max_{i,s} \tau(s)w(i,h) - \eta s w(i,h) \quad \text{s.t. } h = \theta + s,$$

where $\tau(s)$ is the amount of effective working time which is decreasing in s. The parameter governing direct schooling costs η will be used as a supply shifter to endogenously change the levels of skill in the workforce.

Note that although θ and s are additive in producing h, h increases the *log* wage. Thus, ability and schooling are complementary in increasing the wage, and so would be complementary in producing a standard stock of efficiency units of human capital, as measured in Section 2.

Agent θ's first-order conditions are therefore:

$$[s]: \quad \frac{-\tau'(s) + \eta}{\tau(s) - \eta s} = \frac{w_2(i,h)}{w(i,h)}, \tag{13}$$

$$[i]: \quad w_1(i,h) = 0. \tag{14}$$

Substituting (12) into (13) yields:

$$\frac{-\tau'(s) + \eta}{\tau(s) - \eta s} = \frac{a_2(i,h)}{a(i,h)}.$$

This static model of schooling acquisition is an abstraction that yields a simple solution below. Solving a dynamic model would require strong assumptions on the future path of exogenous variables, and would not add greatly to the analysis. First, as with the models considered in Section 2, output, wages, and Mincerian returns can be determined solely from technology and the distribution of different types of labor. That is, the demand for schooling is determined entirely by technology. The supply of skill is determined by (13). At any point, η will be calibrated to match the observed amount of schooling, while $\tau(s)$ will be chosen as a linearization from a lifecycle model, yielding the same *local* elasticity of schooling decisions to relative wages as the lifecycle model.[16]

3.3. Equilibrium

The competitive labor market equilibrium is a set of prices $w(i,h)$, quantities $l(i,h)$, and optimal policy functions for $i(\theta)$ and $s(\theta)$ that satisfy (12)–(14), and market clearing for labor inputs (of different skill level h and in different sectors i).

[16] The one difference is that this would be the elasticity from a change in returns perceived to apply the entire career rather than just one decade. Given the discounting that motivates the calibration of $\tau(s)$ in Section 4, the first decade out of school constitutes about half of the effective career for a person with ten years of schooling. Also, agents may indeed choose their schooling based almost solely on current return, since changes in return may be difficult to forecast.

I solve equivalently for inverse policy mappings of sectors to abilities $\theta(i)$ and sectors to skill levels $h(i)$, which are strictly increasing by Assumptions 1 and 2.[17]

Labor market clearing simplifies to:

$$l(i,h)=\begin{cases} g(\theta(i))\theta'(i)\tau[s(i)] & \text{for } h=\theta(i)+s(i),\\ 0 & \text{otherwise.}\end{cases} \tag{15}$$

In words, the demand for labor of type h working in sector i must equal the supply. For sector-skill combinations that satisfy $h=h(i)$, the supply is the density of workers of the type θ that choose sector i, times the effective time that these workers spend in the labor force given their optimal level of schooling. The $\theta'(i)$ term is the Jacobian term from transforming the density in terms of θ to a density in terms of i. For sector-skill combinations that are not optimal, the supply is zero.

Given the amount of labor, output in sector i is therefore[18]:

$$x(i)=a\big[i,h(i)\big]g\big(\theta(i)\big)\theta'(i)\tau\big[h(i)-\theta(i)\big]. \tag{16}$$

Combining the expression for wages that comes from firm optimization (12) with the household optimality condition in the choice of i (14), one can easily derive the constant elasticity of substitution expression:

$$\frac{a_1(i,h)}{a(i,h)}=\mu\frac{x'(i)}{x(i)}. \tag{17}$$

Taking logs and differentiating (16) and combining with (17) yields a second order differential equation in the matching function $\theta(i)$. Omitting functional dependencies, this equation is:

$$\frac{\theta''}{\theta'}+\left(\frac{g'}{g}-\frac{\tau'}{\tau}\right)\theta'+\left(\frac{a_2}{a}+\frac{\tau'}{\tau}\right)h'+\left(\frac{\mu-1}{\mu}\right)\frac{a_1}{a}=0. \tag{18}$$

This differential equation yields the optimal choice of i given θ. Given the Inada condition on each sector and Assumption 2, the required boundary conditions are that $\theta(\underline{I})=\underline{\theta}$ and $\theta(I)=\overline{\theta}$. The corresponding optimal choice of s (and therefore h) can be easily found by applying (13). These policy functions satisfy the labor market clearing conditions by construction.

3.4. Skill-biased technical change

I model two forms of SBTC.

The first form involves an increase in I, that is, the introduction of new sectors that are more complex and hence more complementary to skilled workers. The increase of I causes worker productivity to increase for two reasons. The first reason is that the complexity level i increases for all (but a measure zero of) workers, and complexity causes the marginal and average productivity of human capital to increase. The second reason is the Romer (1990) growth effect: the range of sectors becomes larger, fewer workers work in each of the more differentiated sectors, and the marginal value of each sector increases. Returning to the wage equation, but substituting $\theta+s$ in for h:

$$w(i,s)=(1-\alpha)A(R)\left[\frac{x(i)}{L}\right]^{-\mu}a(i,\theta+s),$$

an increase in I will lower $[\frac{x(i)}{L}]^{-\mu}$ for all i.

Given this effect, one can also understand the decomposition with respect to industry by discretizing the average wage by i (and realizing the mapping of schooling s to sector i):

$$w_t=\sum_j w_t\big[i_j,s_t(i_j)\big]l_t(i_j)$$

and deriving the terms in (5):

$$\begin{aligned}\frac{\Delta w_t}{\overline{w}_t}&=\frac{\sum_j \overline{w_t}(i_j)\Delta l_t(i_j)}{\overline{w}_t}+\frac{\sum_i \Delta w_t(i_j)\bar{l_t}(i_j)}{\overline{w}_t}\\ &=\frac{\sum_j \overline{w_t}(i_j)\Delta l_t(i_j)}{\overline{w}_t}+(1-\alpha)\frac{\Delta[A(r)(\frac{x(i)}{L})^{-\mu}a_t[i_j,s_t(i_j)]]}{\overline{w}_t}.\end{aligned} \tag{19}$$

[17] In principle, the mapping $s(\theta)$ need not be increasing, but because wages will be increasing in θ, and the calibration matches the positive Mincerian relationship between wages and schooling, $s(\theta)$ is also increasing in the calibrated model.

[18] Eq. (9) assumed that the mass of sectors was distributed across a two-dimensional (h,i) plane. This density would need to be integrated across h in order to reduce the dimensionality to one (the i dimension). The existence of the function $h(i)$ shows that the problem was already one-dimensional, and the mass is distributed along the line $h(i)$. Hence no integration is needed.

Reallocations of workers toward high wage sectors, i.e., $\sum_j \overline{w_t}(i_j)\Delta l_t(i_j)$, will be large when $[\frac{x(i)}{L}]^{-\mu}$ is increasing, so that both the explained wage growth and the residual wage growth will be correlated. I show this numerically in the next section. Furthermore, from this decomposition residual, only the changes in $A(r)$ will be truly "unexplained," since changes in I and γ will be identified in the data.

The second form of SBTC is an increase in the skill-bias within sectors that increases the relative productivity of skilled workers without changing the productivity of the average worker. The interpretation of this second type of technical change as a fall in the relative price of physical capital for skilled workers across all sectors (see (7)) is shown in Appendix B.

The functional form used for the productivity function is:

$$a(i,h) = \exp\left(\frac{i^{1+\phi}h^{1-\phi} + \gamma(h-\bar{h})}{(1+\phi)(1-\phi)}\right).$$

Here ϕ governs the relative weight given to sector and skill in determining an agent's productivity. It also governs the amount of log diminishing returns to skill, with higher ϕ implying more rapidly diminishing returns to skill within a given sector. The parameter γ captures the additional return to skill h that is common to all tasks and governs the within-sector SBTC. Since $\bar{h}$ is the average skill level, γ has no net productivity effect on the average worker. A high γ increases the relative productivity of workers with above average skill, while decreasing the productivity of workers with below average skill.[19]

The derivative below shows that $a(i,h)$ satisfies Assumptions 1–3 and illustrates how both of the skill biases work:

$$\frac{\partial \log a(i,h)}{\partial h} = \frac{1}{(1+\phi)}\left(\frac{i^{1+\phi}}{h^{\phi}} + \frac{\gamma}{1-\phi}\right).$$

Higher I increases the marginal benefit of skill by increasing the i of workers, while higher γ increase the marginal benefit of skill regardless of i.

4. Quantitative results

In this section, I calibrate the model and show comparative statics with respect to η, γ, and I. I then identify the decade-by-decade shifts in these three factors between 1940 and 2000, and quantify the causal contribution of schooling increases.

4.1. Baseline calibration

The parameters to be calibrated are: α, ϕ, μ, $\underline{I}$, I, $g(\theta)$, $\tau(s)$, and η. In practice, ϕ, μ, $\underline{I}$, I, $\tau(s)$, and η are all determined jointly, but I discuss individual targets to illustrate the calibration strategy. I calibrate the model to U.S. data in 1970, the middle of the relevant sample period.

The share of capital α is set to the standard value of 1/3. The remaining technology parameters are ϕ, I_0, and I. The parameter $\mu = 0.65$ implies an elasticity of substitution of 1.5, but this is between tasks. The elasticity between college ($s = 16$) and high school labor ($s = 12$) is somewhat higher, since workers are perfect substitutes within tasks. Katz and Murphy (1992) estimate an the elasticity of substitution between college and high school workers range of 1.4, but this estimate is based on "effective" labor of each type. Using parallel techniques to construct effective labor of those with more than twelve years of schooling and those with up to twelve years and estimating the identical regression on simulated data from the model for the same period (1960–1990), $\mu = 0.65$ yields an elasticity estimate of 1.4.

Together, the remaining technology parameters ϕ (which governs the extent of diminishing returns to skill), $\underline{I}$ and I determine the variation in the relative demand for skill across sectors. Given the decomposition, I choose these three parameters to match three facts: the estimated Mincerian return (6.4%), the standard deviation of wages across the $N = 9$ schooling categories (0.26), and standard deviation of wages across nine grouped industries (0.29) to that of $J = 9$ grouped sectors, evenly spaced between $\underline{I}$ and I. The parameter γ is normalized to zero in 1970. The resulting values for $\underline{I}$ and I are 0.35 and 1.08. The calibrated value of the diminishing return parameter ϕ of 0.76, is identical to the 0.76 value estimated in Kaboski (2007) using cross-country data.

For $g(\theta)$, Cawley et al. (1999) estimate that in log wage terms, the gain from being in a higher ability quartile is about 1.5 times the gain from an additional year of school in the United States. Since the available evidence is on ability quartiles, the uniform distribution of ability θ was used with a range of ability equal to six (i.e. $\bar{\theta} - \underline{\theta} = 4 * 1.5 = 6$).

The remaining parameters govern the supply of schooling. The parameter $\eta = 0$ is calibrated to match the average years of schooling in the economy (11.5 in 1970), and actually equals zero in the baseline. The function $\tau(s)$ is calibrated as $32.8 - s$. This is based on a linearization of discounted effective career length (with schooling beginning at age six and

[19] Manipulation of terms and the redefinition of $\hat{A} \equiv A(R)\exp(\frac{-\gamma\bar{h}}{(1+\phi)(1-\phi)})$ and $\hat{a}(i,h) \equiv \exp(\frac{i^{1+\phi}h^{1-\phi}+\gamma h}{(1+\phi)(1-\phi)})$ shows how skill-biased growth in γ can be viewed as increasing the skill-biased productivity $\hat{a}$, but decreasing the skill-neutral productivity $\hat{A}$.

Table 2
Comparative statics results for various parameter changes causing increase in average schooling

Comparative statics exercise:		Explaining wage increase					Explaining schooling increase			Change in Mincerian return (%)
		Avg. percent wage increase (%)	Schooling decomposition		Sectoral decomposition		Avg. years of schooling increase*	Sectoral decomposition		
			"Explained" increase[a] (%)	Residual (%)	"Between" increase[a] (%)	Residual (%)		"Between" increase[a]	Residual	
		(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
Increase in I	(A)	37	8	30	6	32	0.82	0.61	0.21	+0.2
Increase in γ	(B)	5	8	−3	0	5	0.82	−0.04	0.86	+0.9
Decrease in η	(C)	5	7	−3	0	5	0.82	−0.58	1.40	−0.3

* Parameter values were chosen to produce an increase of 0.82 in average years of schooling, the mean per decade increase experienced by the U.S. workforce (1940–2000). Changes were 0.04 for I (from 1.09 to 1.13), 0.008 for γ (from 0 to 0.008), and −0.048 for η (from 0 to −0.048).

[a] The "explained" increase is the increase explained by changes in quantities of labor inputs at initial wages (or initial schooling levels, respectively). See Eqs. (4)–(6).

retirement at age 65) around an average of 11.5 years of schooling with a discount rate of 4 percent.[20] It implies that for a person with average schooling (11.5), forgone earnings represent about 4.7 ($= 1/(32.8 - 11.5)$) percent of discounted lifetime earnings at that time.

4.2. Comparative statics

I perform comparative statics exercises on the model[21] in which one of the SBTC parameters I or γ, or the supply parameter η are changed from the baseline calibration in order to produce an increase in the average amount of schooling of 0.82 years, the average increase per decade from 1940–2000.[22] In Table 2, I present only the results for changing each parameter separately, since weighted combinations of parameter changes (leading to the same increase in schooling) yield similarly weighted mixes of these results.

Column (i) shows that schooling increases driven by a pure increase in I are associated with large wage increases (37%), while those coming from an increase in γ (within-sector SBTC) or a decrease in η are associated with smaller increases of just (5%). The wage effects of γ are small because changes in γ do not change the productivity of the average worker. The wage effects of η are even lower because of the diminishing returns in the gains from schooling.[23]

Columns (ii)–(iii) give the decompositions of wage growth with respect to schooling (Eq. (4)). The results for the schooling decomposition show that the explained growth is highest for η and γ, which lead to lower wage growth. Thus, the model can potentially explain the puzzling negative correlation between explained growth from schooling decompositions and actual wage growth in Table 1. Since η and γ lead to less actual growth than explained growth, the model can also explain the same phenomenon in the data for the 1970s.

The results for the decomposition of wage growth with respect to sector (Eq. (5)) in columns (iv)–(v) contrast greatly. The explained increases from sectoral decompositions in column (iv) are only substantial when I increases, and so they are highest when wage growth is highest, and could potentially explain the positive correlation with growth.[24]

Column (vi) gives the schooling increase, 0.82 years by construction, while columns (vii)–(viii) give the decomposition of this increase with respect to sector (Eq. (6)). The amount of the schooling increase that can be explained by sectoral reallocations varies greatly and distinguishes the two different forms of SBTC. Increases in I cause shifts of labor toward higher i (and higher s) sectors, and so much of the growth in schooling (0.56 years) is explained by sectoral reallocations. In contrast, increases in γ occur within all sectors, and so they lead to growth within sectors rather than reallocations. Finally, decreases in η actually lead to reallocations toward low schooling sectors.

Finally, column (ix) shows the resulting change in the Mincerian returns, which is crucial for identifying supply driven increases from demand driven increases. Decreases in η lower the Mincerian return, while increases in I and, especially, γ increase the Mincerian return.

[20] A typical lifecycle approach would express discounted lifetime earnings as:

$$\int_s^T e^{-r(t-s_{\text{avg}})} w(s+\theta)\,dt = w(s+\theta)\left[\frac{e^{-r(T-s_{\text{avg}})} - e^{-r(s-s_{\text{avg}})}}{-r}\right],$$

where the discount rate r would encompass the interest rate net of wage growth and a linear return to experience, and s_{avg} is used as the point of reference for discounting since that is the margin between school and work for the average student. Linearizing the bracketed term around s_{avg} yields the $\tau(s)$ that is calibrated.

[21] In these simulations, as in the data, when a baseline wage does not exist for a certain type of labor (e.g., $j = 1$), I use the wage of the next closest type available ($j = 2$). This occurs only for a very small fraction of workers in both the simulations and the data (see Appendix B).

[22] The parameter changes yielding an increase in schooling of 0.82 years: are an increase in I from 1.08 to 1.13; an increase in γ from 0 to 0.0084; and a decrease in η from 0.055 to 0.003.

[23] A second reason this increase is smaller than the Mincerian return is because an upward ability bias in the Mincerian return because $s(\theta)$ is increasing.

[24] Quantitatively, the six-decade correlations for explained growth and actual growth are −0.55 (model) vs. −0.48 (data) for the decomposition with schooling, and 0.99 (model) vs. 0.78 (data: industry) and 0.74 (data: occupation).

Table 3
Policy Relevant Growth Contributions of Schooling, 1940–2000

	Avg. percent wage growth (%) (i)	Total contribution of schooling* (%) (ii)	% of total (iii)	Direct (supply shift) contribution of schooling** (%) (v)	% of total (v)	Indirect (technology induced) contribution of schooling*** (%) (vi)	% of total (ii)
Pre-1970	34	4	12	11	33	−7	−21
Post-1970	10	4	40	−1	−13	5	52
1940–2000	21	5	24	5	22	1	7

* The total effect may not equal the sum of the direct and indirect effects because of rounding error.
** For each decade, the direct effect of schooling is the reduced growth effect of keeping the cost of schooling, η, fixed at its level at the beginning of each decade.
*** For each decade, the total effect of schooling is calculated as the reduced growth had schooling levels remained fixed at their levels at the beginning of each decade. The indirect effect is the total effect minus the direct effect.

4.3. Identification of driving forces

In order to identify the driving forces, I choose the combination of I, γ, and η values to match the growth in average years of schooling, the amount of schooling growth explained by sectoral reallocations, and the change in Mincerian returns in each decade. The rates of change of these parameters differ substantially in the pre-1970 and post-1970 periods.[25] To have a frame of reference, I present the average parameter changes as percentages of the single parameter changes required to induce 0.82 years of schooling in Table 2. In the pre-1970 period, the changes imply schooling increases being driven by an outward supply shift (an average decrease in η by 132 percent of the single parameter change) and a high rate of introduction of new high skill sector (an increase in I by 48 percent), but being negated by technical change within sectors that is biased toward *low* skill workers (a decrease in γ by 46 percent). In contrast, the post-1970 reflects on average an *inward* supply shift (a decrease in η by 24 percent), a slower rate of new sector introduction (an increase in I by just 22 percent), and an increase in within-sector SBTC (an increase in γ by 97 percent). The larger total for the pre-1970 period reflects the slightly larger average increase in years of schooling per decade (0.9 vs. 0.73).

The results of wage growth decompositions on the simulated data compare well, though not perfectly, with the empirical decompositions of Section 2. In particular, the wage growth explained by sector is somewhat less than the growth explained by schooling (simulated: 2 vs. 8 percent, empirically: 3 vs. 5 percent). The wage growth explained by sector is significantly positively correlated with actual growth (simulated: 0.71, empirically: 0.78), while that explained by schooling is negatively correlated with actual wage growth (simulated: −0.09, empirically: −0.48). Finally, residual wage growth has lower volatility in sectoral decompositions than in schooling decompositions (simulated: 13.9 vs. 15.1 percent, empirically: 11.4 vs. 15.2 percent).

More importantly, *unexplained* wage growth—that is, the residual wage growth coming from $A(r)$—is just 6 percentage points in the simulated data compared to 17 percentage points based on the decomposition from the general human capital model in Section 2. Moreover, the standard deviation of residual wage growth is lower in the simulated model (13 vs. 15 percentage points). Hence, modeling the forces behind schooling increases the ability of theory to explain growth, despite the fact that these forces are pinned down independently of growth, and so add no additional free parameters. Finally, these residual $A(r)$ shocks show very little correlation with γ over time (−0.05), which verifies that the assumption that γ has little effect on average productivity is a reasonable one.

4.4. Policy relevant contributions of schooling

I now distinguish two policy relevant contributions of schooling to growth. First, I ask, "Given the shifts in technology in a given decade, how much less (more) growth would have been observed without the decreases (increases) in η?" Heuristically, this "direct contribution" comes from shifting the supply schedule for skilled labor. It is policy relevant because a government policy to promote education, for example, might directly lower η by decreasing the costs of education that students face. The second contribution is the indirect contribution, which answers, "How much additional wage growth would have been lost (gained) had schooling not responded to increases (decreases) in demand driven by technology?" Heuristically, this effect comes from movements *along* the supply schedule for skilled labor. This effect is policy relevant, but only indirectly, since policies such as wage controls, enrollment limits, lack of access to schools, or other frictions could prevent these endogenous responses in schooling.

Table 3 summarizes the average results for each subperiod and for the sample overall. On average, schooling contributed six percentage points (or 29 percent) of total wage growth, with the direct effect contributing 5 percentage points and the

[25] The 1940s and 1970s stand out from other decades in two ways. The 1940s has a rapid outward supply shift but also rapid technical change biased toward low skill workers, which together account for the decline in Mincerian returns. The 1970s is actually characterized by an *outward* shift in supply. These results are consistent with interpretations in the labor literature (see Katz and Murphy, 1992, and Goldin and Margo, 1992).

indirect effect contributing 1 percentage point. The relative importance of these effects varied considerably over the period, however. The direct effects were particularly important in the pre-1970 high growth period when the supply of schooling shifted outward, but actually negative in the post-1970 period as the supply shifted backward (perhaps due to rising costs of tertiary education). In contrast, the indirect effect was substantial in the post-1970 period, contributing 5 percentage points, or 50 percent of overall wage growth. Hence, both direct and indirect effects have played important roles in wage growth, but at different points in time.

5. Conclusion

This paper has developed a theory to understand and quantify the causal forces behind schooling and wage growth. The model here helps distinguish between schooling increases that are distinguished by technological advance—i.e., the growth of new industries and occupations—and high growth rates, and schooling increases driven from either supply shifts or by within-sector growth in the demand for skill. When technological advance is the driving force, educational investments are accompanied by the introduction of new skill-intensive sectors and large movements of labor toward these sectors. The empirical evidence for the United States harmonizes with the model. Schooling increases have been relatively steady, but growth has been greatest in decades when sectoral movements were away from low-skill sectors (such as agriculture or household servants) toward new, high-skill sectors.

The quantitative results suggest that government programs increasing the supply of skill exogenously may be important, but so may policies that enhance the ability of the labor force to respond endogenously to increased demand. On an international level, the model may be important in explaining the varied growth experiences of countries, despite the fact that almost all countries have witnessed substantial increases in education levels.

An avenue for further research would be to either apply the model in a similar fashion to other countries, or to perform the cross-country analog of the time analysis here to see whether sectors are similarly important. Within the context of the model, perhaps the education "miracles" of some countries, such as the East Asian tigers, is not simply their large increases in human capital, but also their structural transformation from agrarian to industrial economies.

Appendix A. Description of data

The data is taken from the IPUMS census data project. Schooling and wage questions are first included in the 1940 census and so limit me to the seven censuses between 1940 and 2000. The variables of interest are annual hours, education, occupation, industry, experience, and wages. Each is discussed in turn.

Annual hours is the product of the number of weeks worked in the previous year and the number of hours per week in the previous week. Both weeks worked and hours per week are coded values.

There are two sources of schooling data: detailed schooling data that varies somewhat across decades and grouped schooling data that is comparable across decades. The detailed data is used for the Mincerian return results and the average education levels in Table 1. Although the answers are top coded, the high codes of 17, 17, 18, 20 and 20 in decades 1940 through 1980 respectively, contain less than 2.6 percent of the total in any given year and so these high codes are not viewed as extremely binding on the data. The 1990 and 2000 years of schooling have been recoded from attainment data. The grouped education categories are: no schooling, 0 to 4 years, 5 to 8 years, 9 years, 10 years, 11 years, 12 years, up to 3 years of college, and four or more years of college. Clearly, these data do not suffer from top-coding.

The standard nine broad occupational classes listed in Table 1 are used to classify occupations. The standard sixteen broad industry categories are grouped into nine categories as described in footnote 10.

The Mincerian return regressions use potential experience calculated in the standard fashion (i.e. age − years of schooling − 6). The alternative decompositions that control for experience used six grouped categories: less than 10, 10 to 19, 20 to 29, 30 to 39, 40 to 49, and over 50 years of experience.

Wages were calculated by dividing annual wage and salary income by annual hours worked. Wages that were either unrealistically high or low were deleted. On the high end, hourly wages above 500 dollars were viewed as unreliable. On the low end wages below fifty percent of the legal minimum wage at the time were dropped from the data set. Finally, real wages were then obtained by deflating with the gross domestic product (GDP) deflator. I use the GDP deflator since I am interested in real product wages rather than real consumption wages.

Finally, a small complication with the decompositions is that certain labor inputs (education-occupation-industry-experience categories) exist in a given decade, but not in the previous decade. These totaled 3.3 percent of the observations across the seven censuses. In these cases, wage data for the previous data must be imputed. The approach taken is to use the average wage of the next level of aggregation—the necessary level of aggregation is determined by the availability of wage data.[26]

[26] The order of aggregation is: (1) aggregate to broader education categories, (2) aggregate to broad industry group, (3) aggregate across all industries, and (4) aggregate across all occupations. No higher level of aggregation was needed.

Appendix B. Heterogeneous capital

The initial production function for firm's as given in Section 2 is:

$$Y = \tilde{A}\left[\int_{\underline{I}}^{\bar{I}} \tilde{x}(i)^{1-\tilde{\mu}}\,di\right]^{\frac{1}{1-\tilde{\mu}}} \quad \text{and} \quad \tilde{\mu} \in (0,1), \tag{B.1}$$

where

$$\tilde{x}(i) = \int_0^\infty \tilde{a}(i,h)k(i,h)^\alpha l(i,h)^{1-\alpha}\,dh. \tag{B.2}$$

The capital and labor decisions of the competitive firm can be analyzed separately. The firm's capital optimization is expressed:

$$\max_{k(i,h)} \tilde{A}\left\{\int_{\underline{I}}^{\bar{I}}\left[\int_0^\infty \tilde{a}(i,h)k(i,h)^\alpha l(i,h)^{1-\alpha}\,dh\right]^{1-\tilde{\mu}} di\right\}^{\frac{1}{1-\tilde{\mu}}} - \tilde{r}(K)\int_{\underline{I}}^{\bar{I}}\int_0^\infty p(h)k(i,h)\,dh\,di,$$

where $\tilde{r}(K)$ is the rental rate of capital and $p(h)$ is the relative cost of capital of skill-level h. Since capital is completely fungible, the rental rate must be constant across all capital and is, therefore, only a function of total capital, $K \equiv \int_{\underline{I}}^{\bar{I}}\int_0^\infty p(h)k(i,h)\,dh\,di$. The first-order condition for this problem is:

$$r(K) = \tilde{A}^{1-\tilde{\mu}}Y^{\tilde{\mu}}\tilde{x}(i)^{\tilde{\mu}}\frac{\tilde{a}(i,h)}{p(h)}\left[\frac{l(i,h)}{k(i,h)}\right]^{1-\alpha} \quad \forall i,h \text{ s.t. } l(i,h) > 0,$$

where $r(K) \equiv \frac{\tilde{r}(K)}{\alpha}$. Using this first-order condition, one can solve for the optimal capital levels:

$$k(i,h) = \left[\frac{\tilde{A}^{1-\tilde{\mu}}Y^{\tilde{\mu}}}{r(K)}\tilde{x}(i)^{-\tilde{\mu}}\frac{\tilde{a}(i,h)}{p(h)}\right]^{\frac{1}{(1-\alpha)}} l(i,h).$$

Substituting this equation into (B.2) yields:

$$\tilde{x}(i) = \int_0^\infty \tilde{a}(i,h)\left[\frac{A^{1-\tilde{\mu}}Y^{\tilde{\mu}}}{r(K)}\tilde{x}(i)^{-\tilde{\mu}}\frac{\tilde{a}(i,h)}{p(h)}\right]^{\frac{\alpha}{(1-\alpha)}} l(i,h)\,dh.$$

Since $\tilde{x}(i)$ is constant with respect to h, it may be factored outside of the integral:

$$\tilde{x}(i) = \left[\tilde{x}(i)^{-\tilde{\mu}}\right]^{\frac{\alpha}{(1-\alpha)}}\int_0^\infty\left[\frac{A^{1-\tilde{\mu}}Y^{\tilde{\mu}}}{r(K)}\right]^{\frac{\alpha}{(1-\alpha)}}\tilde{a}(i,h)^{\frac{1}{(1-\alpha)}}p(h)^{\frac{\alpha}{(1-\alpha)}}l(i,h)\,dh,$$

and $\tilde{x}(i)$ can be isolated on the left-hand side:

$$\tilde{x}(i) = \left\{\int_0^\infty\left[\frac{\tilde{A}^{1-\tilde{\mu}}Y^{\tilde{\mu}}}{r(K)}\right]^{\frac{\alpha}{(1-\alpha)}}\tilde{a}(i,h)^{\frac{1}{(1-\alpha)}}p(h)^{\frac{\alpha}{(1-\alpha)}}l(i,h)\,dh\right\}^{\frac{(1-\alpha)}{1-\alpha+\tilde{\mu}\alpha}}.$$

In addition, since each task is produced using a Cobb–Douglas technology, capital will be paid a constant fraction of product within each task and thus a constant fraction of total final product. Mathematically, $1 = \frac{r(K)K}{Y}$. This expression can be used to substitute for the rental rate. Substituting in and simplifying yields:

$$\tilde{x}(i) = \left[\int_0^\infty\left[\left(\frac{Y}{\tilde{A}}\right)^{\tilde{\mu}-1}K\right]^{\frac{\alpha}{(1-\alpha)}}\tilde{a}(i,h)^{\frac{1}{(1-\alpha)}}p(h)^{\frac{\alpha}{(1-\alpha)}}l(i,h)\,dh\right]^{\frac{(1-\alpha)}{1-\alpha+\tilde{\mu}\alpha}}.$$

Now after substituting this expression for $\tilde{x}(i)$ into (B.1), Y and K can be factored out of the integrals. Isolating Y on the left-hand side and manipulating the exponents yields:

$$\tilde{Y}^{\frac{1}{1-\alpha+\tilde{\mu}\alpha}} = \tilde{A}^{\frac{1}{1-\alpha+\tilde{\mu}\alpha}}K^{\frac{\alpha}{1-\alpha+\tilde{\mu}\alpha}}\left[\int_{\underline{I}}^{I}\left[\int_0^\infty \tilde{a}(i,h)^{\frac{1}{(1-\alpha)}}p(h)^{\frac{\alpha}{(1-\alpha)}}l(i,h)\,dh\right]^{\frac{(1-\alpha)(1-\tilde{\mu})}{1-\alpha+\tilde{\mu}\alpha}} di\right]^{\frac{1}{1-\tilde{\mu}}},$$

$$Y = \tilde{A}K^\alpha\left\{\left[\int_{\underline{I}}^{I}\left[\int_0^\infty \tilde{a}(i,h)^{\frac{1}{(1-\alpha)}}p(h)^{\frac{\alpha}{(1-\alpha)}}l(i,h)\,dh\right]^{\frac{(1-\alpha)(1-\tilde{\mu})}{1-\alpha+\tilde{\mu}\alpha}} di\right]^{\frac{1-\alpha+\tilde{\mu}\alpha}{(1-\tilde{\mu})(1-\alpha)}}\right\}^{1-\alpha}.$$

If I set $p(h)$ equal to a constant one, this equation is identical to (8) and $(9)^{\sim}$ given the definitions of $a(i,h) \equiv \tilde{a}(i,h)^{1/(1-\alpha)}$ and $\mu \equiv \frac{\tilde{\mu}}{1-\alpha+\tilde{\mu}\alpha}$.

If instead I define:

$$p(h) \equiv \left[\exp\left(\gamma(h-\bar{h})/\left[(1+\phi)(1-\phi)\right]\right)\right]^{\frac{(\alpha-1)}{\alpha}},$$
$$\tilde{a}(i,h) \equiv \left[\exp\left(i^{1+\phi}h^{1-\phi}/\left[(1+\phi)(1-\phi)\right]\right)\right]^{(1-\alpha)},$$

it is clear how changes in γ can be viewed as changes in the relative price of capital.

References

Acemoglu, D., 1998. Why do new technologies complement skill? Directed technical change and wage inequality. Quarterly Journal of Economics 113, 1055–1089.
Acemoglu, D., Guerrieri, V., 2005. Capital deepening and non-balanced endogenous growth. Mimeo.
Berman, E., Bound, J., Griliches, Z., 1994. Changes in the demand for skilled labor within U.S. manufacturing industries: Evidence from the annual survey of manufacturers. Quarterly Journal of Economics 109, 367–397.
Berman, E., Bound, J., Machin, S., 1998. Implications of skill-biased technical change: International evidence. Quarterly Journal of Economics 113, 1245–1279.
Bils, M., Klenow, P.J., 2000. Does schooling cause growth? American Economic Review 90, 1160–1183.
Caselli, F., 1999. Technological revolutions. American Economic Review 89, 78–92.
Cawley, J., Heckman, J., Vytlacil, E., 1999. Meritocracy in America: Wages within and across occupations. Industrial Relations 38, 250–296.
Denison, E.F., 1974. Accounting for U.S. Economic Growth, 1929–1969. Brookings Institution, Washington, DC.
Echevarria, C., 1997. Changes in sectoral composition associated with economic growth. International Economic Review 38, 431–452.
Galor, O., Moav, O., 2000. Ability biased technological transition, wage inequality and growth. Quarterly Journal of Economics 115, 469–498.
Galor, O., Tsiddon, D., 1997. Technological progress, mobility and economic growth. American Economic Review 87, 363–382.
Galor, O., Weil, D., 2000. Population, technology, and growth: From Malthusian stagnation to the demographic transition and beyond. American Economic Review 90, 806–828.
Goldin, C., Katz, L., 1998. The origins of capital-skill complementarity. Quarterly Journal of Economics 113, 693–732.
Goldin, C., Margo, R., 1992. The great compression: The wage structure in the United States at mid-century. Quarterly Journal of Economics 107, 1–34.
Gollin, D., Parente, S., Rogerson, R., 2002. The role of agriculture in development. American Economic Review 92, 160–164.
Greenwood, J., Yorukoglu, M., 1997. 1974. Carnegie–Rochester Conference Series on Public Policy 46, 49–95.
Hall, R.E., Jones, C.I., 1999. Why do some countries produce so much more output per worker than others? Quarterly Journal of Economics 114, 83–116.
Hendricks, L.A., 2002. How important is human capital for development? Evidence from immigrant earnings. American Economic Review 92, 198–219.
Jorgenson, D.W., Gollop, F.M., Fraumeni, B., 1987. Productivity and U.S. Economic Growth. Harvard University Press, Cambridge, MA.
Juhn, C., Murphy, K.M., Pierce, B., 1993. Wage inequality and the rise in returns to skill. Journal of Political Economy 101, 410–442.
Kaboski, J.P., 2007. Endogenous Mincerian returns: Explaining returns to schooling and output across countries. Mimeo.
Katz, L., Murphy, K.M., 1992. Changes in the wage structure 1963–1987: Supply and demand factors. Quarterly Journal of Economics 107, 35–78.
Klenow, P.J., Rodriguez-Clare, A., 1997. The neoclassical revival in growth economics: Has it gone too far? In: Bernanke, B.S., Rotemberg, J.J. (Eds.), NBER Macroeconomics Annual 1997. MIT Press, Cambridge, MA.
Kongsamut, P., Rebelo, S., Xie, D., 2001. Beyond balanced growth. Review of Economic Studies 68, 869–882.
Krueger, A., Lindahl, M., 2001. Education for growth: Why and for whom? Journal of Economic Literature 39, 1101–1136.
Krusell, P., Ohanian, L.E., Rios-Rull, J.V., Violante, G.L., 2000. Capital-skill complementarity and inequality: A macroeconomic analysis. Econometrica 68, 1029–1053.
Lucas Jr., R.E., 1988. On the mechanics of economic development. Journal of Monetary Economics XXII, 3–42.
Machin, S., Van Reenen, J., 1998. Technology and changes in skill structure: Evidence from seven OECD countries. Quarterly Journal of Economics 113, 1215–1244.
Mankiw, G., Romer, D., Weil, D., 1992. A contribution to the empirics of economic growth. Quarterly Journal of Economics 107, 407–437.
Manuelli, R., Seshadri, A., 2005. Human Capital and the Wealth of Nations.
Matsuyama, K., 1992. A simple model of sectoral adjustments. Review of Economic Studies 59, 375–387.
Matsuyama, K., 2002. The rise of mass consumption societies. Journal of Political Economy 110, 1035–1070.
Mincer, J., 1974. Schooling, Experience and Earnings. Columbia University Press, New York.
Murphy, K.M., Welch, F., 1993. Occupational change and the demand for skill, 1940–1990. American Economic Review 83, 122–126.
Ngai, L.R., Pissarides, C., 2004. Structural change in multi-sector model of growth. Mimeo.
Romer, P.M., 1990. Endogenous technological change. Journal of Political Economy 98, S71–S102.
Xiang, C., 2005. New goods and the relative demand for skilled labor. Review of Economics and Statistics, 285–298.
Young, A., 1995. The tyranny of numbers: Confronting the statistical realities of the East Asian growth experience. Quarterly Journal of Economics 110, 641–680.