



Measuring Aggregate Human Capital

CASEY B. MULLIGAN

The University of Chicago

XAVIER SALA-I-MARTIN

Department of Economics, Columbia University, 420 West 118 Street, New York, NY 10027 and Universitat Pompeu Fabra

We construct a set of human capital indexes for the states of the United States for each census year starting in 1940. To do so, we propose a new methodology for the construction of index numbers in panel data sets. Our method is based on an optimal approach by which we choose the “best” set of index numbers by minimizing the expected estimation error subject to some research constraints. Some of the empirical findings are that the stock of human capital in the United States grew twice as rapidly as the average years of schooling and that human capital inequality across states went up during the 1980s (while the dispersion of schooling actually fell). We conclude that using the average years of schooling for the empirical study of existing growth models may be misleading.

Keywords: human capital, Index numbers, Divisia index, multilateral comparisons, traveling salesman problem

JEL classification: C43, C82, O49

1. Introduction

The recent growth literature has emphasized the importance of education and human capital in the process of economic growth and development. Following Lucas (1988), a number of authors have constructed multisectoral growth models to study the dynamic behavior of the economy when agents face the decision to invest in their own bodies as well as in physical assets.

The behavior of these models cannot be estimated and tested empirically until good measures of the stock of human capital become available. This article proposes a new methodology for the construction of human capital stocks as well as a set of estimates of such stocks for the states of the United States for each census year starting in 1940. We follow the theoretical literature by assuming, throughout the article, that the production function of an economy can be expressed as a stable function of aggregate inputs. One of these inputs is what we call *human capital*, the input associated with the labor force. Because, on one hand, workers are heterogeneous and have different productivities but, on the other hand, the theoretical models consider labor aggregates, the first stage of an empirical study must aggregate the heterogeneous workers.¹ We will maintain the assumption that there exists a stable aggregator function whose functional form is unknown, except that it is homogeneous of degree one. Our problem is to generate estimates of the stocks of human capital when

the functional form of the true aggregator is unknown.

The early empirical economists used school enrollment rates as a proxy for the stock of human capital (see, for example, Barro, 1991). Kyriakou (1992) attempts to estimate average years of schooling by using a benchmark year and the estimates of school enrollments. Barro and Lee (1993, 1996) compute educational attainment distributions for a large cross-section of countries for the period 1960 to 1990.² These attainment distributions allow them to construct a measure of average years of schooling, which is then used in the literature as the best measure of human capital stocks.

The average years of schooling is not necessarily a good measure of human capital for a variety of reasons. First, it assumes that workers of each education category are perfect substitutes for workers of all other categories. Second, it assumes that the productivity differentials among workers with different levels of education are proportional to their years of schooling. For example, it assumes that a worker with 16 years of schooling is 16 times as productive as a worker with one year of schooling, regardless of their wage rate differentials. Third, the elasticity of substitution across workers of different groups is assumed to be constant always and everywhere. Fourth, one year of schooling is assumed to deliver the same increase in skill always and everywhere. In particular, it does so regardless of the field of study and the quality of the teachers and the education infrastructure.

In a previous paper (Mulligan and Sala-i-Martin, 1997), we attempted to solve these problems by constructing a measure of human capital based on labor income. Our intuition was that the quality of a person would be related to the wage rate he receives in the marketplace. If the type of education a person received was very useful, the markets would reward him with a high wage. Similarly, if a person devoted himself to study a field that was not very useful from a production point of view (such as sixteenth-century moral reasoning or certain types of theoretical economics), then the productive human capital of that person would be low, and his wage rate would indicate so.

The key point was that a worker's wage would also depend on the amount of other aggregate inputs available in the economy in which he works. For example, wages in 1940 were lower than in 1990, not only because skills were lower but also because the amount of physical capital and the level of technology were lower. To eliminate the effect of the aggregate variables from the individual's wage rate, we divided each person's wage rate by the wage rate of the zero-skill worker (who, by definition, has no skill). Therefore, our measure of human capital for an economy was the weighted sum of all workers, where the weights were the ratio of their wage to the wage of the zero-human-capital worker. This is equivalent to the aggregate wage bill divided by the wage of the zero-schooling worker.

This measure, which we called *labor-income-based* (LIB) measure of human capital, had the advantage of being consistent with variable elasticities of substitution across types of workers. Furthermore, it did not impose all workers with the same amount of education to have the same amount of skill (so if they had studied different things, their productivities were allowed to differ). It also allowed for changes in the relative productivities over time and across economies.

The main problem with the LIB measure was that we had to assume that the zero-schooling worker (which we took as the unskilled worker) had the same amount of skill always and everywhere to be able to use him as a numeraire. We also had to assume that this worker was

a perfect substitute for all the others (although we allowed for any degree of substitutability among all the other types). To the extent that these assumptions do not hold in the data, our measures are biased (and some will argue strongly against these assumptions.) A second problem was that when the relative wages among workers changed for reasons other than technological shocks, then our measures would unrealistically reflect movements in the stock of human capital. In cases when these prices change substantially, our LIB measure would tend to fluctuate unnecessarily. In this article we propose a different measure of human capital based on the education attainment of the labor force. We think that it resolves the shortcomings of the previous one.

Along with human capital estimates, we present a new methodology for the construction of *index numbers* that allow for cross-sectional as well as time-series comparisons. Hence, this article represents a contribution to the index number literature. Unlike our predecessors, we propose an *optimal* index number in that we choose the index numbers that minimize a function of the expected error made when human capital indexes are compared across economies. As Diewert (1987) points out, the authors of the existing index number literature are unable to choose among the scores of functional forms proposed over the course of a century. The best they can do is to check whether a particular index number passes a number of desirable “tests” (for example, one may want to check that an index number is invariant to changes in scale, that it is monotonic and transitive, and so on; see Diewert, 1987, for an excellent exposition of which numbers satisfy which properties). Two problems arise, however: which tests are appropriate, and which number do we pick among the ones that pass the desirable tests?

This article presents a solution to the problem of choosing among index numbers by suggesting that the optimal index number is the one that minimizes an expected-error criterion. We implement this idea by constructing optimal human-capital indexes across the states of the United States for all census years starting in 1940. Before we do so, in Section 2 we construct a human-capital index by estimating the parameters of a translogarithmic aggregator. In Sections 3 and 4 we describe Divisia indexes over time and across economies. Section 5 describes the spaces of configurations over which we will search for the optimal index number. In Section 6 we propose an *expected-error* or *cost* criterion. The minimization of this criterion will deliver the optimal index numbers. Section 7 compares the *costs* of the various index number schemes. Section 8 describes the behavior of our favorite human-capital indexes across states and over time. The final section concludes. We should mention that we actually construct all the index numbers we describe throughout the article, even when we reject their validity. Some descriptive statistics for all of them are reported in the appendix tables.

2. The Translog Aggregator: A Regression Approach

Let X be the vector of educational attainment numbers, x_{ia} (where x_{ia} is the fraction of category i workers in economy a).³ We assume that the vector of inputs can be aggregated into a human-capital number using the aggregator function $F[X]$. The number $H = F[x_{1a}, \dots, x_{Na}]$ is what we call the stock of human capital in economy A . As we mentioned in the introduction, our goal is to use an aggregator function $F(\cdot)$ to compute the stock of

human capital in a panel of economies in a way that allows for both intertemporal and cross-sectional comparisons. Although the functional form $F(\cdot)$ is unknown in general, in this section we assume that it is known to be translogarithmic:

$$\ln H = \ln F(X) = \alpha'(\ln X) + \frac{1}{2}(\ln X)'\beta(\ln X), \quad (1)$$

where H is the vector of aggregate human capital numbers for M economies, $H = [H_1, \dots, H_M]$, β is a symmetric matrix of elasticities. For now, the functional form is assumed to be known—although its parameters, α and β , are not. The assumption of linear homogeneity implied by the assumption of a translog $F(\cdot)$ requires that the columns of α must add up to 1 and the column sums of β must equal zero. Differentiating (1) with respect to $(\ln X)$, we get the following relationship:

$$v = \alpha + \beta(\ln X), \quad (2)$$

where v is the vector of income elasticities. Under the usual assumptions about perfect competition, these elasticities correspond to the *share of labor income* earned by the workers of category i . Since we have data on all of these shares as well as on the attainment distribution $([x_{1a}, \dots, x_{Na}])$, the parameters α and β in equation (2) could be estimated using available data. Once the parameters α and β in (2) were estimated, we could use (1) to estimate the level of human capital in each economy.

The translog production function is a second-order Taylor approximation to any production function (see Diewert, 1976). Hence, the method just described is likely to work well when the range of inputs is close to the approximation point or the true production function is not too different from quadratic. However, when we estimate human-capital stocks across a vast range of economies over a long period of time (as it will be the case in the present article where we will attempt to compute the stocks of human capital across 48 U.S. states from 1940 to 1990), it is unlikely that we will always be “close” to the point around which we choose to approximate. A second argument against using this estimation approach is that there is no reason to believe that the aggregator function is quadratic (in the log space). Hence, the regression method for estimating translog aggregators is likely to generate inaccurate estimates of the levels of human capital.

Even though we did not think that this estimation method would deliver good estimates of the stock of human capital, we estimated the system of equations in (2) for the states of the United States for the period 1940 to 1990. We used U.S. Census data on six educational attainment categories broken down by gender for each of the six decades and for the 48 contiguous states. The data are described in Appendix B. A brief analysis of the results is reported in Appendix A.

The translog specification has strong predictions for the relationship between the production elasticities and the stocks of inputs: the relation is exactly linear as in (2) and the coefficients in β are symmetric. Using the estimates for the states of the United States reported in Appendix A, we tested the restrictions imposed by the translog specification and easily rejected them. Hence, perhaps the translog function is not a very good description of the process that generates human capital data in the real world, at least in the United States for the aforementioned period.

3. Divisia Index Over Time

To compute the rate of change of the aggregate stock of human capital between periods t and $t + 1$ for a typical economy, the Divisia index approach suggests the following calculation (see Törnquist, 1936, and Diewert, 1976):

$$\log \frac{H_{t+1}}{H_t} \equiv \sum_{i=1}^N \left(\frac{v_{i,t} + v_{i,t+1}}{2} \right) \log \frac{x_{i,t+1}}{x_{i,t}}, \quad v_{i,t} \equiv \frac{\partial \log F}{\partial \log x_i} \Big|_{x=x_t}, \quad (3)$$

where $v_{i,t}$ is the i th category period t share of labor income. Note that, by using the average share of the two periods, the Divisia index is the average of the Laspeyres index (which uses the shares at the beginning of the period, $v_{i,t}$; see Laspeyres, 1871) and the Paasche index (which uses the shares at the end of the period, $v_{i,t+1}$; see Paasche, 1874).

Because it makes bilateral comparisons, the Divisia index formula generates a series of estimates of the rates of change of the aggregate stock of human capital for each economy A. If we normalize the initial stock to one, the time series of growth rates can be used to compute a time series of levels of human capital. In other words, by establishing a series of bilateral comparisons between successive periods, we can generate a time series of growth rates and levels of human capital for each economy, which allows for multilateral comparisons of the stocks of human capital of one economy over time. We can, in principle, generate one of such series for each economy. Notice, however, that each of the time series will be normalized to one in the first period so the cross-sectional comparison among them will be meaningless. Thus, this method cannot be used to establish cross-sectional comparisons.

One common justification for the use of the Laspeyres or Paasche indexes is that they are a first-order approximation to the rate of change of *any* aggregator. The Divisia index estimate of the rate of change because it is just the simple average of the Laspeyres and Paasche rate of change indices, is a second-order approximation to the rate of change of *any* aggregator.⁴ A key difference between the Laspeyres, Paasche, or Divisia indexes and the regression method outlined in the previous section is that the former approaches *continuously update* the approximation points. For example, when we compute the growth rate of the economy between 1940 and 1941 using the Laspeyres index, we make a Taylor approximation around 1940, but when we do the same for the period 1989 to 1990 we approximate around 1989. It follows that the Laspeyres, Paasche, or Divisia approximations will be valid, even if we attempt to compute human capital stocks over a vast range of economies, as long as any two *successive* economies are not “too far apart” so that the higher-order terms of the Taylor expansion series can be safely ignored (the concept of distance between two economies will be made more precise in Section 6).

4. Cross-Sectional Divisia Index

To compare the stocks of human capital of economies A and B, we could compute the *rate of change of H* between A and B (where A and B are different regions from a geographical

point of view, rather than the same economy at two different points in time) by applying a cross-sectional version of (3):

$$\log \frac{H_B}{H_A} \equiv \sum_{i=1}^N \left(\frac{v_{i,a} + v_{i,b}}{2} \right) \log \frac{x_{i,B}}{x_{i,A}}, \quad v_{i,A} \equiv \left. \frac{\partial \log F}{\partial \log x_i} \right|_{X=X_A}. \quad (4)$$

If we normalize the level of human capital of the *first* economy to one, then the rate of change between the first and *second* economies generates the level for the *second* economy. The rate of change between the *second* and *third* delivers the level for the *third* economy and so on. Repeating this procedure for all the economies we can construct a full cross-section of levels of human capital. In other words, by applying bilateral cross-sectional comparisons between successive pairs of economies, we will arrive at a system of multilateral cross-sectional comparisons.

The key question is: which economy is the *first* one? More generally, what's the order in which the successive growth rates should be computed? In a time series, it seems natural (almost unquestionable) to assume that 1941 should come after 1940 and before 1942. But what is the *correct* order of economies when we have a cross-section? Perhaps more important, does the ordering matter?⁵ It turns out that the actual ordering matters significantly (see Section 7). Hence, close attention needs to be paid to the way the economies are ordered.

One way to solve the problem is to avoid ordering altogether. The trick is to pick a base economy in a given year and then calculate the rate of change of human capital between each economy and the base economy. By normalizing the base economy's stock of human capital to one, this procedure generates a cross-section of measures of human capital that allows for multilateral comparisons. Kravis (1984) called this the *star* method because the base economy plays a starring role: each region is compared to it and it alone.

If we want to compute a panel of human-capital stocks (that is, an index number for each economy at each point in time), we could simply use the same methodology by neglecting the fact that different economies exist at different points in time. In other words, we could pick California in 1960 (CA60) as the base economy. We could then compute the rate of change between each of the 48 economies for each of the years and CA60. This would generate one human-capital number for each state and year. We call this the *superstar* method because the base economy is the star not only for its own time period but also for all periods: all economies of all times are compared to it and to it along.⁶

Of course, the problem with the *star system* is its lack of invariance to the choice of the base (star or superstar) economy. Different choices for the base economy will give rise to different indexes. Fisher (1922) recognized that the simplest way of achieving symmetry was to average base-specific indexes over all possible bases. We estimated a human-capital index using this *average star* and *average superstar* methods. We report some comparative summary statistics in Appendix A. Another possibility is to try to find the *optimal base* economy. After defining optimality in Section 6, we will compute human-capital indexes using optimal basis in Section 7.

A second way to deal with the ordering problem is to choose what we call the *optimal order*. The key to the optimal order is to remember that the Divisia index can be thought of as a continuously updating second-order approximation to any aggregator function $F(X)$. We know that the error we make when we use a Taylor approximation is *small* when we are

close to the point around which we approximate. Therefore, to minimize the error in the computation of our human-capital index, we need to remain as close as possible to the point of approximation. But if we want to make sure that we do, we need to define a measure of distance between vectors of inputs. The optimal ordering will be the one that minimizes some function of such a distance. According to this method, the economy that should be *next* to economy A is the one whose inputs (that is, whose vector of attainment numbers) are the *closest* to those in economy A. The next two steps will be to define a measure of distance and a space over which we will minimize such distance.

Before we proceed, we would like to make an important point. The creation of a multi-lateral index of the sort proposed here suggests that the best way to compare the indexes for two economies (that is, the best way to make a bilateral comparison) is to use data on other economies. For example, when comparing the 1950 stocks of human capital in Oklahoma and Utah, it might be good to use information on economies who are “in between” these two states. In other words, although the comparison between Oklahoma and Utah could be done by direct application of the bilateral index formula in equation (4), it might be better to compare Oklahoma with its “next” economy and then compare this “next” economy with the next and so on until we reach Utah. The intuition for this important result should by now be familiar: if the vectors of educational attainments in Oklahoma and Utah are very different, then the error we make by a direct quadratic approximation might be very large. In fact, it might be much larger than the sum of errors of all the intermediate comparisons.

5. The Choice of Space

5.1. Ω : The Unilink Space

Define *configuration* as a collection of bilateral links between pairs of economies. In the previous section, we suggested a few ways to link economies. One configuration is the one often used in the construction of time-series indexes—a linear arrangement of economies. For example, the 1940 economy comes first, the 1940 economy is linked to the 1941 economy, the 1941 economy is linked to the 1942 economy, etc.

A single economy acts as the base economy (that is, the economy to which all other economies are directly linked) in another configuration—the star system of Kravis (1984). In both setups, a bilateral index formula such as Divisia is used to compute the growth rate of the index across links. However, neither the linear arrangement nor the Kravis star is the only conceivable way to link economies. This section defines some possible arrangements.

Since we want to come up with *unique* measures of human capital, we will consider only spaces (sets of configurations) that link any two economies through a *single* possible route.⁷ Define the set of all such configurations of M economies to be the “ M -noded *unilink* space.” We use the notation $\Omega(M)$ to denote the M -noded *unilink* space.

A *unilink* space, $\Omega(4)$, can be illustrated in an example. Figure 1a shows each of the elements of the unilink space: each of the four economies can be directly linked to as many as $M - 1$ (three in this example) economies, and all economies are linked (at least indirectly). However, for a configuration to be an element of the *unilink* space, there must be only one chain of bilateral comparisons between any two economies. In other words,

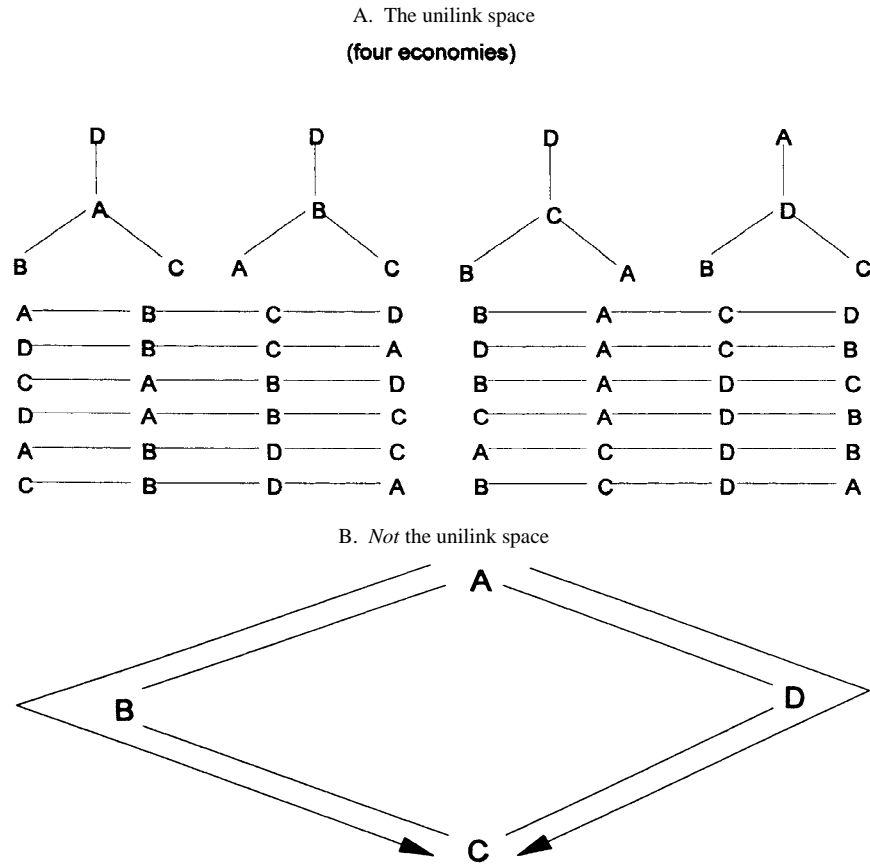


Figure 1. The unilink space and *not* the unilink space (four economies).

the “route” linking any two economies must be unique. For example, the configuration displayed in Figure 1b is not an element of the *unilink* space because economies *A* and *C* can be compared in two ways: (1) a bilateral comparison *A* and *B* together with a bilateral comparison *B* and *C* and (2) a bilateral comparison *A* and *D* together with a bilateral comparison *D* and *C*.

Although the *unilink* space rules out some configurations, there are still clearly many ways to arrange economies so that a system of multilateral comparisons can be constructed using the Divisia bilateral index. Our goal is to search in this very large space to find the optimal configuration.⁸ However, the space is so large that we need to economize on search effort by considering particular subspaces of the unilink space.⁹ The remainder of this section therefore defines some subspaces.

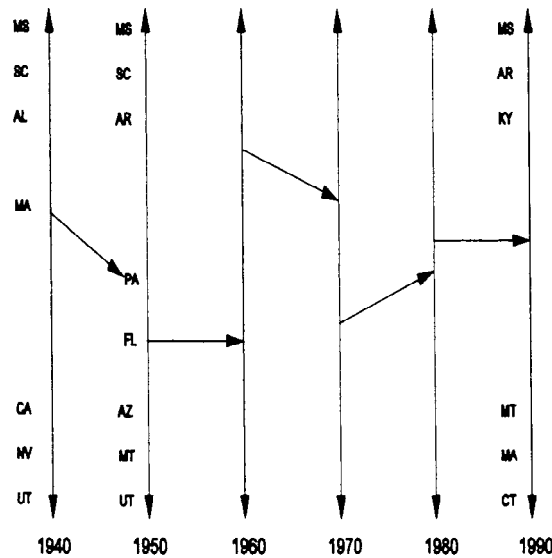


Figure 2. Geographical traveling salesman.

5.2. Ω_1 : *Geographical Traveling Salesman (GTS)*

The first subspace is what we call the *geographic traveling salesman (GTS)*, as we will think of a salesman who has to go to each of the economies *in a particular year*. He can choose the ordering as well as the economy in which he starts his trip. His only constraint is that he has to stop in every economy once and only once. For every cross-section, there are $48!$ possible configurations.

Elements of the GTS space can differ in the way that economies are ordered with a cross-section. Another way that elements of the GTS space can differ is in the way that the different cross-sections are linked. For example, one configuration may always link period t Connecticut with period $t + 1$ Alabama. Another one may link all six Mississippi economies in chronological order. An example of an element of the GTS space is depicted in Figure 2.

When we implement our methodology using the GTS space, we will minimize a cost function by searching over elements or configurations of the GTS space. This computational problem is a variation of what mathematicians call the traveling salesman problem. Mathematicians and computer scientists have been trying to find the solution to the traveling salesman problem for years (see Press et al., 1988, and Garey and Johnson, 1979). Although an exact solution has not been discovered, a number of computational methods have been found to yield close approximations (see Kirkpatrick et al., 1983). In this article we use a version of the simulated annealing method reported in Press et al. (1988). For the 48 economies in any one of our example's cross-sections, the algorithm requires minutes of a Pentium III's computer time to (approximately) minimize a cost function over $\Omega_1(48)$ —as

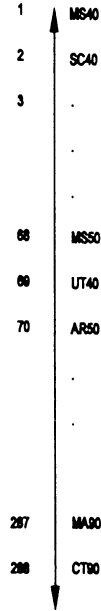


Figure 3. Intertemporal traveling salesman.

opposed to the thousands of years required to approximately minimize a cost function over the entire unilink space $\Omega(48)$.

5.3. Ω_2 : The Intertemporal Traveling Salesman (ITS)

The GTS space allows the salesman to travel across all regions in one particular year but does not allow him to travel over time. For example, he is not allowed to go from CT40 to OK50 and back to MA40. The *intertemporal traveling salesman* space allows for these intertemporal trips. In principle there is no reason not to allow for these intertemporal trips: it could very well be the case that WI60 is very “similar” (where “similar” will be defined by a distance metric) to AL70 and MA50. If this is the case, we would like to allow for these economies to be “next” to each other. Since we have 48 states and 6 time periods, we have a total of 288 economies. We allow the traveling salesman to visit any two consecutive economies of his choosing, regardless of the time period in which these two economies happen to be (see Figure 3). We also allow the traveling salesman to choose the initial point of his trip. Since we have now 288 economies, for each configuration, the cost function contains $288 \times 287/2$ elements. There are $288!$ configurations.

Notice that the GTS space Ω_1 is not a subset of the ITS space Ω_2 because elements of ITS must link the *entire panel* in a linear fashion. Elements of GTS must link each cross-section in a linear fashion, but the separate cross-sections can be linked in arbitrary ways (as long

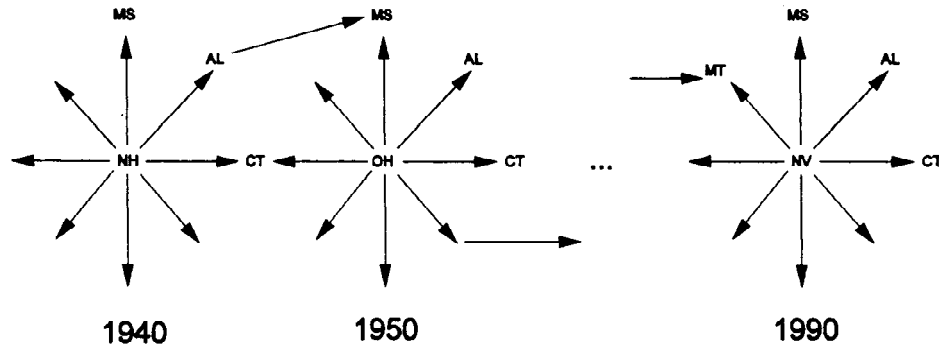


Figure 4. Geographical star system.

as the configuration is part of the unilink space).

5.4. Ω_3 : The Star Space

This space follows the *star* method described in Section 4. Different elements of the star space Ω_3 can differ for two reasons. First, they may have different economies play the starring role in each cross-section. For example, one configuration has Oklahoma playing the starring role in every cross-section. Another element has New York as the star. Yet another configuration has Illinois as the 1940 star, Georgia as the 1950 star, Washington as the 1960 star, California as the 1970 star, West Virginia as the 1980 star, and Florida as the 1990 star.

A second difference across elements of the star space is the way in which they link the cross-sections. For example, one configuration may always link period t Connecticut with period $t + 1$ Alabama. Another one may link all six Mississippi economies in chronological order. A typical configuration of this subspace is depicted in Figure 4.

An interesting point is that the famous Kravis's system utilized by the Summers and Heston (1991) data set of multilateral GDP comparisons across nations uses a version of this geographical star system. In a few words, Summers and Heston pick a star economy that the authors think is more or less in the middle of all the economies. This economy plays the starring role. A bilateral Divisia comparison is then applied between each economy and the star. This allows the authors to generate a cross-sectional series of GDP index numbers.¹⁰

5.5. Ω_4 : The Superstar Space

This space is based on what we called the *superstar system* in Section 3. For all of the 288 cross-section-time-series economies, pick a base economy that will play the role of *superstar*. For example, in Figure 5 the superstar economy is Ohio in 1960 (OH60).

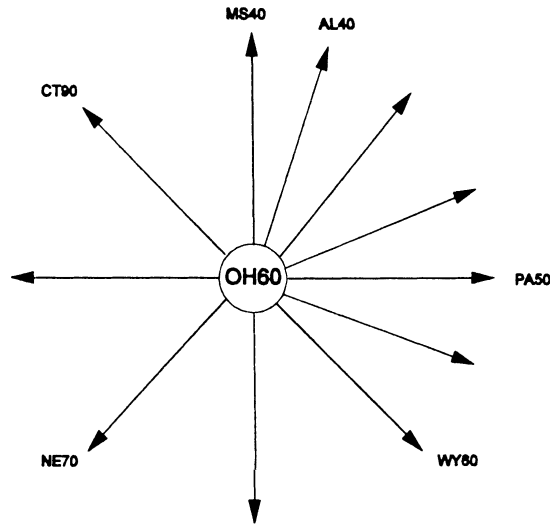


Figure 5. The superstar system.

Different elements of the superstar space Ω_4 differ in only one way: the choice of the starring economy. Hence, $\Omega_4(M)$ has M elements.

6. Measures of Distance Between Vectors of Inputs: A Cost-Function Approach

Our basic method has three steps:

1. Arrange the economies (choose a configuration or element ω of the unilink space Ω).
2. Beginning with a numeraire economy, use a bilateral index number formula to compute the human capital of each of the other economies.
3. Repeat steps 1 and 2 until the *best* configuration is found.

This section defines *best*.

Our previous discussion suggests a definition of *best*. We argue that the bilateral index number formulas are most appropriate for two economies that are “close together” in the input space. One possible definition of distance between two linked economies A and B is

the weighted sum of squares of the differences between category i inputs for economies A and B :

$$\sum_{i=1}^N \left(\frac{v_{i,A} + v_{i,B}}{2} \right) \left(\log \frac{x_{i,B}}{x_{i,A}} \right)^2, \quad (5)$$

where the weights are the average for the two economies (as defined in (4)). A strategy might be to minimize a sum of distances of type (5).

Note that this definition of distance has some appealing properties. For example, it suggests that if two economies have identical attainment distributions (so that $x_{i,A} = x_{i,B}$ for all i), then their distance is zero. The optimal ordering will therefore have to place these two economies right next to each other. If two economies differ only in the attainment of categories for which the shares $v_{i,A}$ and $v_{i,B}$ are small (that is, categories that appear to be economically “unimportant”), then the distance between these two economies is still small. In other words, differences in unimportant categories are counted little. On the other hand, this definition assigns a large distance, when the differences between A and B in the important categories are large.

Although reasonable, this definition of distance is arbitrary.¹¹ An alternative and possibly better approach would entail the minimization of the error we expect to make when we compute the human capital index. In other words, since we do not know the true aggregator function, any bilateral comparison based on an index number that we construct will be in error. For example, we might report that economy A ’s index is 10 percent larger than economy B ’s, when in fact economy A ’s is 2 percent smaller.

Let $\hat{H}_A(\omega)$ and $\hat{H}_B(\omega)$ be the estimated indexes of human capital derived from configuration ω for economies A and B , respectively. The true relative size of economies A and B is given by H_A and H_B . Hence, the error we make when we evaluate the relative sizes is given by

$$d_{A,B}(\omega) \equiv \left(\log \frac{\hat{H}_A(\omega)}{\hat{H}_B(\omega)} - \log \frac{H_A}{H_B} \right)^2. \quad (6)$$

Our notation includes ω to emphasize that, because the configuration determines how the human-capital indexes are estimated, the error $d_{A,B}$ depends on the configuration.

Our multilateral index will allow us to compare any two pairs of economies with the comparison between each pair involving an error term like (6). A measure of the quality of bilateral comparisons resulting from a set of index numbers is given by the average of the squared bilateral errors:

$$\frac{1}{M^2} \sum_{i=1}^M \sum_{j=1}^M d_{i,j}(\omega). \quad (7)$$

Of course, by definition of the index number problem, one does not know the value of each of the $d_{i,j}$ terms. We proceed by computing expectations of equation (7) with respect to a set of priors $\Pi(\omega)$ about the magnitude of the approximation errors $d_{i,j}$. The resulting

expectation $C(\omega)$ is the cost of the configuration ω :

$$C(\omega) \equiv E \left(\frac{1}{M^2} \sum_{i=1}^M \sum_{j=1}^M d_{i,j}(\omega) \middle| \Pi(\omega) \right). \quad (8)$$

Our notation allows for the possibility that priors about estimation errors may depend on the configuration.

The *optimal configuration* is the element of the *unilink* space that minimizes cost:

$$\omega^* \equiv \operatorname{argmin}_{\omega \in \Omega} C(\omega). \quad (9)$$

Assuming that the expectations are well defined, a minimum clearly exists because the unilink space is finite. The optimal multilateral system of index numbers is computed using a bilateral index number formula—such as Divisia—and the optimal configuration.

At this point, the optimal configuration is not computable for two reasons. First, priors about the magnitude of bilateral errors are unspecified. Second, the unilink space is a very large space—too large to search exhaustively. In Appendix C, we suggest five approximations that allow us to arrive at an estimate of the optimal configuration.

If the true aggregator function were translog, using all of the data to estimate the translog parameters would be the best way to construct index numbers for a cross-section, time-series, or a panel. Use of a chained Divisia index (or some other strategy) is therefore *an admission that the true aggregator function is not translog*. True, the Divisia is exact when the true aggregator is translog, but each link of the Divisia chain is computed using data from only two economies. The justification for neglecting the data for the other $(M - 2)$ economies must be that the true aggregator function is not translog. Our approach has been to be explicit about our doubts about the exactness of the translog approximation and to show that *those doubts imply that there is an optimal strategy* for computing index numbers and that the optimal strategy is not necessarily the commonly used Divisia chain applied to an ad-hoc ordering of economies.

We propose to use a Divisia chain but to choose the optimal configuration of economies. We seek to minimize the expected bilateral comparison error and show how that criteria can be computed as a function of observables. We propose a set of configurations, the unilink space Ω , over which the criteria should be minimized.

With the current computer technology, exact solutions to our minimization problem cannot be found. However, we propose a feasible algorithm for approximating a solution to the problem. Section 7 below shows that, despite the inexactness of our solution, we can substantially improve on commonly used index number strategies.

Note that the cost function we have derived in this section is similar in spirit to the concept of distance we postulated in (5): it involves the power sums of differences in education categories. Each of the sums is weighted by the average of the shares (the distance measure (5) uses the shares directly, the cost function uses the square of the shares). The main difference between the two concepts is that the cost function (10) involves 6th-power terms (rather than only quadratic terms) as well as cross-products. It turns out that use of the cost (10) instead of the measure of distance (5) matters little. To document this point, in Appendix A, Appendix Table 5, we report the correlation between what will end up being

Table 1. Costs of computing human capital indexes.

A. Optimal measures

| | Space | Pseudocost | Cost | Relative to GTS |
|----------------------------------|------------|------------|-------|-----------------|
| Geographical traveling salesman | Ω_1 | 675.3 | 14.30 | 1 |
| Intertemporal traveling salesman | Ω_2 | 399.3 | 61.16 | 4.3 |
| Star | Ω_3 | 11908 | > 42 | > 2.9 |
| Superstar (OK70 is superstar) | Ω_4 | 476670 | 3299 | 230.7 |

B. Nonoptimal measures

| | Pseudocost | Relative to GTS |
|---------------------------------|------------|-----------------|
| Star system (average) | 93679.00 | 142.50 |
| Superstar system (average) | 4418800.00 | 6716.60 |
| 1940 average years of schooling | 27735.00 | 42.20 |
| Alphabetical | 116600.00 | 177.40 |

Notes: There are 12 categories (six levels of schooling times two genders). Optimal numbers are constructed by minimizing the cost function over the relevant space. Nonoptimal function involve no minimization. Pseudocost is the simple sum of the cost of each link of configuration (weighting each link equally). Cost is the average squared bilateral comparison error. For the space $\Omega_4(M)$, the two differ by the factor $2(M - 1)/M^2$.

our favorite measure of human capital, computed using the cost function derived in this section and a similar measure using the ad-hoc cost function postulated in (5). For 1940, 1960, 1980, and 1990, the correlation is above 0.94. For 1950, it is slightly under 0.9 and for 1970 it is 0.84. If we use all decades at the same time, the correlation is over 0.97. Hence, it turns out that the exact cost function used for estimating human capital indexes does not matter a great deal. What turns out to be more important is to search beyond some of the simple configurations that have been considered in the index number literature.

7. Costs of Each of the Methods and Spaces

The first panel of Table 1 reports the costs and pseudocosts associated with each of the methods outlined in the previous sections. Among the measures we estimate using our optimal methods (that is, the measures that involve the minimization of a cost criterion over some space), we find the best results when we search over the *geographical traveling salesman* space (GTS). The lowest pseudocost that we found in the GTS space was 657.3. The actual cost of the pseudocost minimizing configuration turned out to be 14.3. The second best appears to be the *intertemporal traveling salesman* (ITS) with a cost of 61.2. We found what we thought to be a good configuration from the ITS space by searching with the pseudocost criterion; the best pseudocost that we could achieve on this space was 399.3. The worst space was the *superstar* with an cost of 3299 (which is 230.7 times as expensive as the GTS). The star space delivered an intermediate result with a cost of more than 42 (more than 2.9 times the cost of GTS).¹² These results suggest that the best panel of human capital numbers is the one generated by the GTS system.

An interesting finding from Table 1 is that Kravis's *star methodology* is very costly. If the star was to be chosen optimally, then the overall cost would be more than 42 (more than 2.9 times the cost of GTS). Normally, the base economy is not chosen optimally so the cost of the method used by Kravis and colleagues is likely to be higher than that. Remember that the Summers and Heston data set uses this methodology to make international comparisons of levels of income, output, consumption, and investment. Our calculations suggest that their methodology is inefficient from the point of view of minimizing the sum of expected errors.

Panel B of Table 1 reports the pseudocosts associated with four measures based on nonoptimal orderings (that is, measures that do not involve the minimization of the cost function over some space). The first one is the *average star system* described in Section 4. Remember that this method involves choosing a base or star economy for each period and computing the bilateral comparison between each economy within the same period and the base economy. A cross-section of indexes is computed. The procedure is repeated 48 times (each economy is picked as base economy once), and the average of all 48 cross-sections is taken as the best index of human capital. We then linked each period through the minimum cost link. The pseudocost associated with this method is 93,679, which is 142.5 times more than the pseudocost of our optimal measure.

The second row reports the pseudocost associated with the *average superstar system*. This method is similar to the average star system, except that for each base economy we compute 287 bilateral comparisons corresponding to each of the 48 states for each of the six census years. This generates a 288×1 vector of stocks human capital. We repeat the procedure by using each of the 288 economies as the base economy and then average the results. The total pseudocost of this method is 4,418,800, which is 6,716 times as expensive as our optimal method with the GTS space.

Finally, we report the pseudocost of using two random orderings. The first is the pseudocost of ordering states by 1940 *average years of schooling*, which is estimated to be 27,735 (42 times more costly than GTS.) The second is *alphabetical* an ordering. Its pseudocost is 116,600 (177 times the pseudocost of GTS).

We conclude that the best measure is the geographical traveling salesman (GTS), in the sense that it minimizes the sum of expected errors as defined in (8) or (10). In Table 2 we decompose the total pseudocost of GTS by year and by link. The 1940 cross section pseudocosts 19.27 and the 1950 cross-section pseudocosts 48.08. An interesting aspect of this table is the pseudocost associated with the 1970 cross-section: 388.20, that is, more than one half of the overall pseudocost of computing GTS comes from the 1970 cross-section. We believe that these pseudocosts are related to the confidence we should put in each of the measures (that is, the cost is related to the variance of the measurement error.) Hence, when we use the human capital data derived in this section, we should keep in mind that the error with which we measure the 1970 data is larger than the error with which we measure all other decades.

Finally, Table 2 also reports the pseudocosts associated with each of the intertemporal links. It turns out that the intertemporal links are the following: MA40 to PA50 (pseudocost = 0.0010), FL50 to MS60 (pseudocost = 0.0024), FL60 to LA70 (pseudocost = 0.0065), CA70 to AL80 (pseudocost = 0.0113), and CA80 to TX90 (pseudocost = 0.0967.)

Table 2. Pseudocost decomposition of GTS by year.

| Year | Pseudocost |
|------|------------|
| 1940 | 19.30 |
| 1950 | 48.10 |
| 1960 | 5.90 |
| 1970 | 388.20 |
| 1980 | 39.20 |
| 1990 | 156.20 |

| Intertemporal Links | Pseudocost |
|---------------------|------------|
| MA40-PA50 | 0.0010 |
| FL50-MS60 | 0.0024 |
| FL60-LA70 | 0.0065 |
| CA70-AL80 | 0.0113 |
| CA80-TX90 | 0.0967 |
| Total | 675.3 |

8. Results: A Measure of Human Capital

As indicated in the previous section, our method of analysis suggests that the geographical traveling salesman (GTS) yields the lowest expected error. We will therefore adopt the GTS as our measure of human capital for the rest of the article. Table 3 reports the logarithm of the computed panel of human capital indexes. Each of the indexes is expressed in California 1980 units of human capital. Hence, a negative number for economy i means that the level of human for that economy is less than the level in California 1980 so that the log is negative.

Figure 6 displays the behavior of over time of the log of our measure of human capital averaged by census region (the exact figures are reported in the last four rows of Table 3). We note that in 1940 the region with the highest stock of human capital was the West, followed by the Midwest, Northeast, and the South. The dispersion of human capital across the four census regions was quite high. By 1990, the West had lost its leading position to the Northeast, while the South was still last. We can see in the figure that four regions were much closer together in 1990 than they were in 1940.

A couple of interesting facts are worth highlighting. First of all, the average stock of human capital for the West fell between 1940 and 1950. The reason for this is that during the 1940s the West experienced a rapid increase in the participation of females with low schooling. Second, we note that the relative positions of the four regions remained constant between 1940 and 1970. During the 1970s the Northeast took the second position from the Midwest, and during the 1980s it took the leading position from the West. During the 1980s we see that the growth rate for the Northeast was substantially higher than that of the remaining regions of the country. This could be the reflection of the much publicized migration of skilled people toward New England during the 1980s.

Table 3. Human capital by state and year.

| | State | 1940 | 1950 | 1960 | 1970 | 1980 | 1990 |
|----|-------|--------|--------|--------|--------|--------|--------|
| 1 | AL | -0.390 | -0.270 | -0.151 | -0.098 | -0.050 | -0.018 |
| 2 | AZ | -0.151 | -0.120 | -0.076 | -0.035 | -0.011 | 0.014 |
| 3 | AR | -0.343 | -0.290 | -0.151 | -0.117 | -0.068 | -0.041 |
| 4 | CA | -0.090 | -0.133 | -0.050 | -0.031 | 0.000 | 0.032 |
| 5 | CO | -0.119 | -0.138 | -0.040 | 0.036 | 0.003 | 0.037 |
| 6 | CT | -0.207 | -0.158 | -0.101 | -0.045 | -0.002 | 0.053 |
| 7 | DE | -0.206 | -0.164 | -0.090 | -0.075 | -0.002 | -0.006 |
| 8 | FL | -0.250 | -0.163 | -0.107 | -0.065 | -0.037 | -0.010 |
| 9 | GA | -0.388 | -0.280 | -0.166 | -0.113 | -0.052 | -0.011 |
| 10 | ID | -0.107 | -0.131 | -0.051 | -0.072 | -0.002 | -0.017 |
| 11 | IL | -0.190 | -0.152 | -0.094 | -0.076 | -0.022 | 0.023 |
| 12 | IN | -0.161 | -0.148 | -0.092 | -0.083 | -0.049 | -0.019 |
| 13 | IA | -0.155 | -0.149 | -0.086 | -0.008 | -0.046 | -0.019 |
| 14 | KS | -0.126 | -0.123 | -0.048 | 0.017 | -0.023 | 0.018 |
| 15 | KY | -0.309 | -0.229 | -0.149 | -0.118 | -0.069 | -0.039 |
| 16 | LA | -0.359 | -0.250 | -0.141 | -0.097 | -0.036 | -0.016 |
| 17 | ME | -0.190 | -0.181 | -0.103 | -0.086 | -0.037 | 0.001 |
| 18 | MD | -0.242 | -0.155 | -0.099 | -0.080 | -0.011 | 0.037 |
| 19 | MA | -0.174 | -0.149 | -0.095 | -0.049 | 0.001 | 0.053 |
| 20 | MI | -0.162 | -0.149 | -0.093 | -0.071 | -0.020 | -0.004 |
| 21 | MN | -0.188 | -0.165 | -0.083 | -0.005 | -0.022 | 0.016 |
| 22 | MS | -0.431 | -0.332 | -0.178 | -0.114 | -0.065 | -0.052 |
| 23 | MO | -0.221 | -0.168 | -0.109 | -0.015 | -0.041 | -0.011 |
| 24 | MT | -0.128 | -0.100 | -0.059 | 0.046 | -0.008 | 0.013 |
| 25 | NE | -0.147 | -0.150 | -0.069 | -0.092 | -0.019 | 0.009 |
| 26 | NV | -0.078 | -0.160 | -0.085 | -0.061 | -0.029 | -0.031 |
| 27 | NH | -0.198 | -0.162 | -0.113 | -0.091 | -0.043 | 0.006 |
| 28 | NJ | -0.210 | -0.153 | -0.100 | -0.051 | -0.007 | 0.039 |
| 29 | NM | -0.215 | -0.165 | -0.044 | -0.021 | -0.019 | 0.021 |
| 30 | NY | -0.184 | -0.152 | -0.097 | -0.044 | -0.010 | 0.031 |
| 31 | NC | -0.331 | -0.277 | -0.168 | -0.137 | -0.070 | -0.030 |
| 32 | ND | -0.226 | -0.179 | -0.126 | -0.025 | -0.036 | -0.016 |
| 33 | OH | -0.153 | -0.151 | -0.087 | -0.076 | -0.026 | -0.004 |
| 34 | OK | -0.191 | -0.160 | -0.073 | -0.060 | -0.016 | -0.013 |
| 35 | OR | -0.104 | -0.144 | -0.053 | -0.050 | -0.004 | 0.014 |
| 36 | PA | -0.205 | -0.167 | -0.102 | -0.083 | -0.036 | 0.006 |
| 37 | RI | -0.233 | -0.172 | -0.128 | -0.123 | -0.047 | -0.005 |
| 38 | SC | -0.406 | -0.306 | -0.180 | -0.147 | -0.072 | -0.032 |
| 39 | SD | -0.202 | -0.156 | -0.123 | -0.072 | -0.043 | -0.039 |
| 40 | TN | -0.297 | -0.231 | -0.162 | -0.127 | -0.061 | -0.025 |
| 41 | TX | -0.194 | -0.159 | -0.097 | -0.080 | -0.025 | 0.008 |
| 42 | UT | -0.072 | -0.098 | -0.020 | 0.043 | 0.009 | 0.020 |
| 43 | VT | -0.189 | -0.170 | -0.101 | -0.098 | -0.038 | -0.003 |
| 44 | VA | -0.280 | -0.179 | -0.118 | -0.065 | -0.016 | 0.028 |
| 45 | WA | -0.098 | -0.123 | -0.038 | 0.015 | -0.004 | 0.025 |
| 46 | WV | -0.222 | -0.205 | -0.097 | -0.092 | -0.041 | -0.030 |
| 47 | WI | -0.206 | -0.167 | -0.106 | -0.083 | -0.039 | -0.007 |
| 48 | WY | -0.119 | -0.120 | -0.049 | -0.109 | 0.001 | 0.016 |

(continued)

Table 3. (continued).

| | | | | | | |
|-----------|--------|--------|--------|--------|--------|--------|
| US. | -0.209 | -0.175 | -0.099 | -0.064 | -0.028 | 0.000 |
| s.d. | 0.088 | 0.053 | 0.039 | 0.046 | 0.022 | 0.025 |
| Northeast | -0.199 | -0.163 | -0.105 | -0.075 | -0.024 | 0.020 |
| South | -0.302 | -0.228 | -0.133 | -0.099 | -0.043 | -0.016 |
| Midwest | -0.178 | -0.155 | -0.093 | -0.049 | -0.032 | -0.004 |
| West | -0.116 | -0.130 | -0.051 | -0.022 | -0.006 | 0.013 |

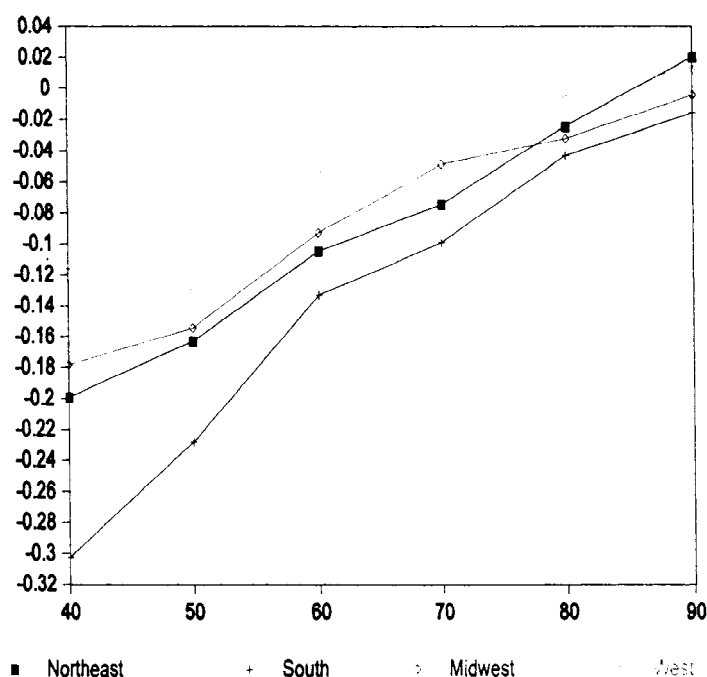


Figure 6. Human capital by census region.

Figure 7 summarizes the behavior of the average stock of human capital for each of the 48 states between 1940 and 1990. The horizontal axis has the 48 states ranked by increasing levels of human capital in 1940. We see that the economy with the lowest stock of human capital was Mississippi, followed South Carolina, Alabama, and Georgia. At the upper end, we have Utah, Nevada, California, and Washington (a complete yearly ranking of states is reported in Table 4).

The line marked with squares corresponds to the level of human capital in 1940, and by construction it is monotonic and upward sloping. The line corresponding line for 1950 is represented by daggers. We note three characteristics. First, the 1950s data are no longer monotonic. This indicates that the relative ranking has changed (some states have improved

Table 4. State rankings by year.

| | 1940 | 1950 | 1960 | 1970 | 1980 | 1990 |
|----|------|------|------|------|------|------|
| 1 | UT | UT | UT | MT | UT | MA |
| 2 | NV | MT | WA | UT | CO | CT |
| 3 | CA | WY | CO | CO | MA | NJ |
| 4 | WA | AZ | NM | KS | WY | CO |
| 5 | OR | WA | KS | WA | CA-> | MD |
| 6 | ID | KS | WY | MN | CT | CA |
| 7 | CO | ID | CA | IA | ID | NY |
| 8 | WY | CA | ID | MO | DE | VA |
| 9 | KS | CO | OR | NM | OR | WA |
| 10 | MT | OR | MT | ND | WA | IL |
| 11 | NE | IN | NE | CA-> | NJ | NM |
| 12 | AZ | MI | OK | AZ | MT | UT |
| 13 | OH | MA | AZ | NY | NY | KS |
| 14 | IA | IA | MN | CT | MD | MN |
| 15 | IN | NE | NV | MA | AZ | WY |
| 16 | MI | OH | IA | OR | OK | OR |
| 17 | MA-> | IL | OH | NJ | VA | AZ |
| 18 | NY | NY | DE | OK | NM | MT |
| 19 | MN | NJ | IN | NV | NE | NE |
| 20 | VT | MD | MI | VA | MI | ->TX |
| 21 | IL | SD | IL | FL | MN | PA |
| 22 | ME | CT | MA | MI | IL | NH |
| 23 | OK | TX | NY | SD | KS | ME |
| 24 | TX | OK | WV | ID | TX | VT |
| 25 | NH | NV | TX | DE | OH | OH |
| 26 | SD | NH | MD | OH | NV | MI |
| 27 | PA | FL-> | NJ | IL | LA | RI |
| 28 | WI | DE | VT | MD | ND | DE |
| 29 | DE | MN | CT | TX | ->PA | WI |
| 30 | CT | NM | PA | PA | FL | FL |
| 31 | NJ | WI | ME | IN | ME | MO |
| 32 | NM | ->PA | WI | WI | VT | GA |
| 33 | MO | MO | FL-> | ME | WI | OK |
| 34 | WV | VT | MO | NH | WV | LA |
| 35 | ND | RI | NH | NE | MO | ND |
| 36 | RI | VA | VA | WV | SD | ID |
| 37 | MD | ND | SD | ->LA | NH | AL |
| 38 | FL | ME | ND | AL | IA | IN |
| 39 | VA | WV | RI | VT | RI | IA |
| 40 | TN | KY | LA | WY | IN | TN |
| 41 | KY | TN | KY | GA | AL | WV |
| 42 | NC | LA | AL | MS | GA | NC |
| 43 | AR | AL | AR | AR | TN | NV |
| 44 | LA | NC | TN | KY | MS | SC |
| 45 | GA | GA | GA | RI | AR | KY |
| 46 | AL | AR | NC | TN | KY | SD |
| 47 | SC | SC | ->MS | NC | NC | AR |
| 48 | MS | MS | SC | SC | SC | MS |

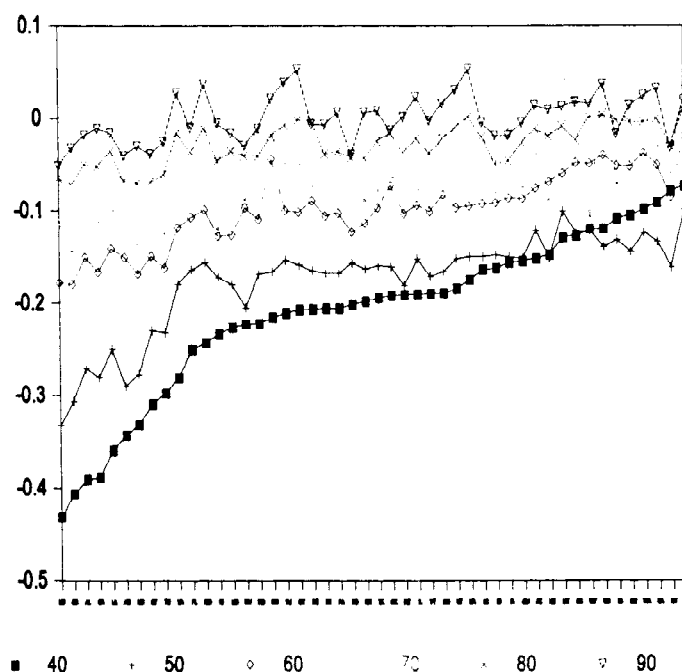


Figure 7. Human capital across states and over time.

their relative position, and as a result other states have worsened). Second, the overall trend seems to be less steep than the line for 1940. This indicates that the states that used to have less human capital have experienced a higher growth rate in the stock of human capital. The same phenomena can be observed for the stocks in 1960, 1970, 1980, and 1990. In particular, the line for 1990 is quite flat and not monotonic. This means that the ranking in 1940 is very different from the ranking in 1990, and the variance in 1990 is smaller than the variance in 1940. The states with the largest stocks in 1990 were Massachusetts, Connecticut, and New Jersey, while the states with the least amounts of human capital in 1990 were Mississippi, Arkansas, and South Dakota.

It is interesting to analyze the speed at which our measures of human capital regress to the mean. Table 5 reports some of the results. Note that this speed of convergence is higher than the estimated speed of convergence of income per capita, which is close to 2 percent per year (see, for example, Barro and Sala-i-Martin, 1992). The striking negative relation between the initial level and the subsequent growth rate in the stock of human capital can also be appreciated in Figure 8.

The rest of the rows of Table 5 suggest that the speed of convergence h was far from constant across decades. After two decades of large speeds of convergence ($= 0.0755$ (s.e. $= 0.0057$) for the 1940s and $= 0.0446$ (s.e. $= 0.0078$) in the 1950s), the speed of convergence achieved a relatively low value in the 1960s (0.0121 (s.e. $= 0.0173$)). The largest

Table 5. Convergence of human capital.

| Period | β (s.e.) | R^2 [s.e.] |
|------------|-------------------|------------------|
| 1940–1950 | 0.039 (0.005) | 0.92 [0.0004] |
| 1940–1950 | 0.058 (0.006) | 0.80 [0.0020] |
| 1950–1960 | 0.045 (0.008) | 0.52 [0.0019] |
| 1960–1970 | 0.012 (0.013) | 0.02 [0.0032] |
| 1970–1980 | 0.118 (0.017) | 0.78 [0.0017] |
| 1980–1990 | 0.009 (0.011) | 0.02 [0.0016] |
| Restricted | 0.055 (0.004) | — — |

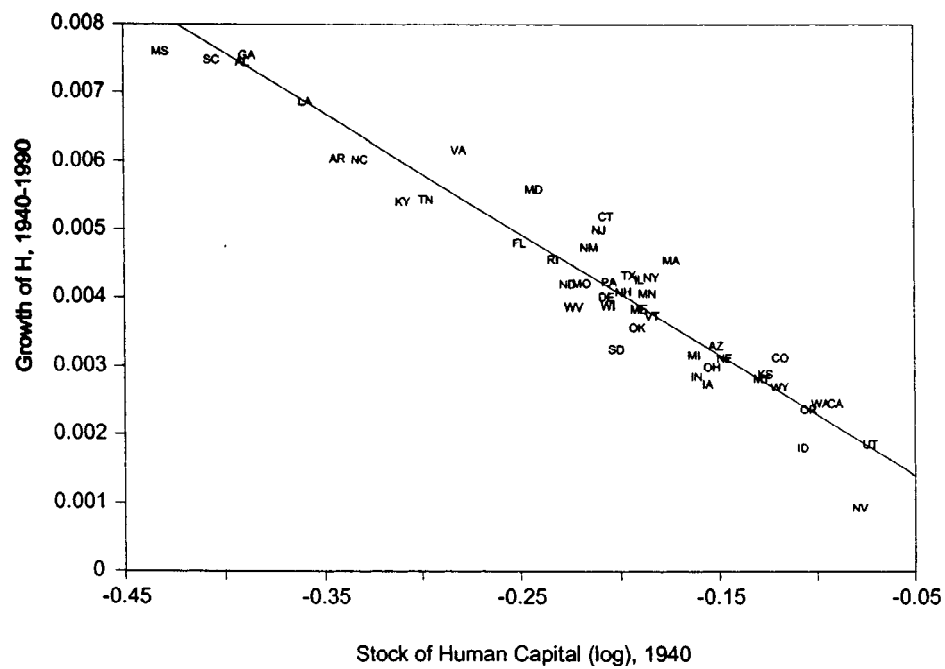


Figure 8. Convergence of human capital, 1940–1990.

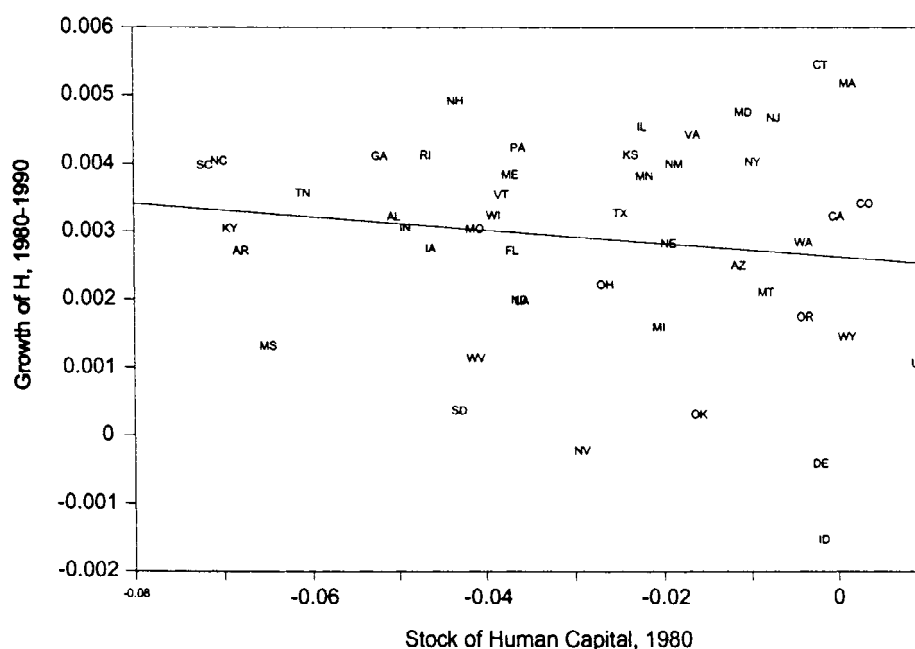


Figure 9. Convergence of human capital, 1980–1990.

speed of convergence observe in our sample occurs in the 1970s (0.1176, s.e. = 0.0173). We note that the process of convergence halted during the 1980s (the estimated speed 0.0093(s.e. = 0.0109)). The no-relation between growth and the initial stock of human capital during the 1980s can be also appreciated in Figure 9. We note in this figure that the stocks of human capital for the states in the Northeast (Massachusetts, Connecticut, New Jersey, Maryland, and New York) experience the largest growth rates during the 1980s, despite the fact that the corresponding stocks in 1980 were not small. However, we also see in the figure that even if we abstract from the northeastern states, the negative relation is far from obvious. In other words, even though it is true that the stock of human capital in the Northeast experienced an extraordinary increase during the 1980s, the story behind the lack of convergence during this same period cannot be solely explained by the Northeastern experience (this is especially interesting because a lot of economic observers have suggested that “strange” behavior of the state economies during the 1980s can be explained by the enormous increase in the stock of human capital in the Northeast—and especially New England).

The last row of Table 5 reports the estimated speed of convergence when all subperiods are restricted to have the same coefficients. The estimated speed is 0.0554(s.e. = 0.0043). It is interesting to point out that this speed is substantially larger than the speed of convergence of conventional measures output such as personal income or gross state product.¹³

Figure 10 analyzes the behavior of the dispersion of the stock of human capital over time. We note that the dispersion in 1940 was quite large: the standard deviation of the

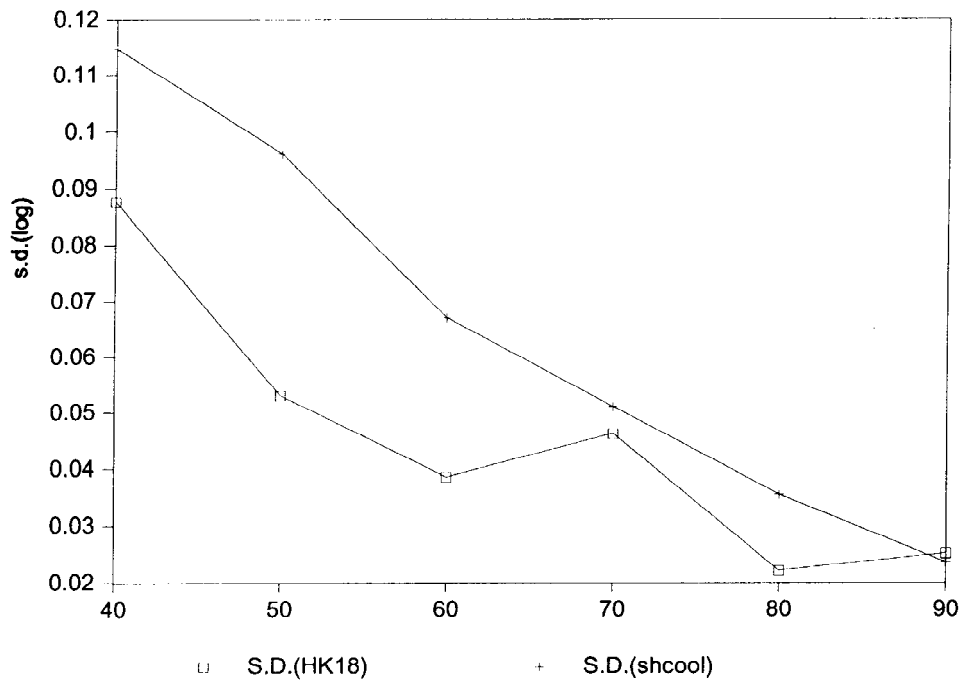


Figure 10. Dispersion of human capital and schooling over time.

log of H was 0.09. After a dramatic fall during the 1940 and 1950s, dispersion increased in the 1960s. The process of convergence resumed between 1970 and 1980. Dispersion increased during the 1980s. This increase in dispersion is often associated with the increase in income inequality which occurred during the Reagan years. It is interesting to compare the behavior of human capital with the behavior of the average years of schooling. Figure 10 also shows the dispersion of the average years of schooling over time. Notice that it falls monotonically. Hence, a researcher who uses schooling as a proxy for human capital will tend to get misleading results.

9. Comparing the Various Measures Described in the Article

In Appendix A we present summary statistics for all the measures that have been described in the article. We report estimates of optimal measures (that is, measures that correspond to the minimization of the cost function over some subspace of the unilink space) and nonoptimal measures (which correspond to some ad-hoc criteria). Among the optimal measures, we consider the optimal estimates when we searched for configurations in the following spaces: *geographical traveling salesman*, Ω_1 , *intertemporal traveling salesman*, Ω_2 , *star*, Ω_3 , and *superstar*, Ω_4 . Among the nonoptimal methods, we report the estimates for *star system*,

superstar system, estimation of the *translog* production function, the *labor-income-based* measure of Mulligan and Sala-i-Martin (1997) and the *average years of schooling*.

Three basic messages arise from the appendix. First, if one wants to use the optimal approach to computing cross-sectional index numbers, the exact ordering matters. We can see this by comparing the optimal estimates using the various orderings with our best measure, the GTS. It is important, therefore, to find the optimal order by minimizing the expected error. Second, if one decides to use nonoptimal methods, then one's estimates will greatly vary with the method adopted. This is the classical index number problem described by Diewert (1987): how do we choose among alternative index numbers? The third lesson is that, for the particular application of this article (namely, the construction of human capital indexes across the states of the United States), the set of estimates delivered by the star space ends up being very similar to our optimal measure. Hence, it is very possible that the breakdown of economies in yearly stars might deliver good enough estimates in other applications. This result is interesting because the star system is much easier to implement than the geographical traveling salesman (remember, again, that the star methodology is the one employed by Summers and Heston to create their cross-country estimates of GDP indexes). A corollary of all this is that if the econometrician is good at picking the star or base, then the index number estimates generated by the star method may not be too bad.

10. Conclusion

The main goal of this article is the construction of a panel data set of human capital indexes for the states of the United States for the period 1940 to 1990. We find that slightly different procedures for constructing human capital indexes yield very different sets of numbers. It is for this reason that we propose a new methodology for the selection of a group of index numbers. As economists, we can't resist thinking that the best way to estimate index numbers (or the best way to do anything in life, for that matter) is to optimize. Hence, we propose to use a Divisia chain, but to choose the optimal configuration of economies, we seek to minimize the expected bilateral comparison error and show how that criteria can be computed as a function of observables. We propose a set of configurations, the unilink space Ω , over which the criteria should be minimized.

With the current computer technology, exact solutions to our minimization problem cannot be found. However, we propose a feasible algorithm for approximating a solution to the problem. Section 7 shows that, despite the inexactness of our solution, we can substantially improve on commonly used index number strategies.

Using this new methodology, we arrive at the optimal estimates of the stock of human capital across the states of the United States. Our empirical analysis of these estimates suggests some interesting results. For example, we find that the stock of human capital grew twice as rapidly as the average years of schooling (the measure that is usually taken to be a good proxy for human capital in the growth literature). We also find that the dispersion of the stock of human capital across the United States increased during the 1980s. This is an interesting finding because the dispersion of the average years of schooling decreased during the same period. Hence, an economist using the average years of schooling as the estimate of the stock of human capital may wrongly conclude that the increase in income

inequality occurred during the 1980s had nothing to do with the process of human capital accumulation.

Finally, we find that Kravis's *star methodology* used by Summers and Heston (1991) to create multilateral comparisons of GDP across countries was very costly in the sense of generating a high sum of expected errors. However, for the states of the United States, the actual estimates delivered by the optimal measure and the star method are very similar, which suggests that if one picks the star (or base) economy carefully, then the index number estimates generated by the star method may not be too bad. A next research step is to use the optimal methodology proposed in this article to construct human capital indexes for a panel of countries to be used along with the Summers and Heston data set.

Appendix A. A Brief Analysis of the Various Measures

In this appendix we compare the results of estimating human capital using the various methods and spaces described throughout the article. We report the following measures.

1. Minimum cost over the geographical traveling salesman space (described in Sections 5 and 6 in the text). This measure is called GTS. As discussed in the text, this is our favorite measure.
2. Minimum cost over the intertemporal traveling salesman space (described in Sections 5 and 6 in the text). Called ITS.
3. Minimum cost over the star space (described in Sections 5 and 6 in the text). Called STAR.
4. Minimum cost over the superstar space (described in Sections 5 and 6 in the text). Called SUPERSTAR.
5. Nonoptimal star system (described in Section 4 in the text). Called STAR (AV).
6. Nonoptimal superstar system (described in Section 4 in the text). Called SUPER (AV).
7. Regression estimates of translog production function coefficients (described in Section 2 in the text). Called TRANSLOG.

We compare these seven measures with

8. Labor-income-based human capital measure reported by Mulligan and Sala-i-Martin (1997) (a brief description of this method appears in Section 1). Called LIB.
9. Average years of schooling. Called SCHOOL.

To economize on space, we do not report each of the estimates of human capital for each state and year for each of the measures. Instead, we report some interesting summary statistics.

The first two panels of Appendix Table 1 reports the behavior of the U.S. averages according to each of the nine measures. The first panel reports the (logarithm of the) levels

Appendix Table 1. Behavior of U.S. aggregates according to the various measures.

| Logs | 1940 | 1950 | 1960 | 1970 | 1980 | 1990 |
|----------|--------|--------|--------|--------|--------|--------|
| GTS | -0.209 | -0.175 | -0.099 | -0.064 | -0.028 | 0.000 |
| ITS | -0.297 | -0.195 | -0.120 | -0.118 | -0.069 | -0.119 |
| STAR | -0.181 | -0.129 | -0.064 | -0.060 | -0.022 | 0.006 |
| SPERSTAR | -0.179 | -0.101 | -0.052 | -0.018 | 0.035 | 0.066 |
| STAR(AV) | -0.202 | -0.152 | -0.079 | -0.065 | -0.023 | 0.004 |
| SPER(AV) | -0.153 | -0.088 | -0.049 | -0.033 | -0.011 | -0.010 |
| TRANSLOG | -0.990 | -0.790 | -0.574 | -0.357 | -0.133 | 0.014 |
| LIB | 3.701 | 3.383 | 3.687 | 3.662 | 3.871 | 4.393 |
| SCHOOL | 2.160 | 2.251 | 2.360 | 2.429 | 2.533 | 2.576 |

| Growth Rates | 1940–1950 | 1950–1960 | 1960–1970 | 1970–1980 | 1980–1990 | 1940–1990 |
|--------------|-----------|-----------|-----------|-----------|-----------|-----------|
| GTS | 0.0034 | 0.0076 | 0.0035 | 0.0036 | 0.0029 | 0.0042 |
| ITS | 0.0102 | 0.0074 | 0.0002 | 0.0049 | -0.0049 | 0.0036 |
| STAR | 0.0052 | 0.0064 | 0.0005 | 0.0037 | 0.0029 | 0.0037 |
| SPERSTAR | 0.0078 | 0.0049 | 0.0034 | 0.0052 | 0.0031 | 0.0049 |
| STAR(AV) | 0.0049 | 0.0074 | 0.0013 | 0.0043 | 0.0026 | 0.0041 |
| SPER(AV) | 0.0065 | 0.0039 | 0.0016 | 0.0023 | 0.0001 | 0.0029 |
| TRANSLOG | 0.0200 | 0.0216 | 0.0217 | 0.0224 | 0.0147 | 0.0201 |
| LIB | -0.0318 | 0.0303 | -0.0025 | 0.0209 | 0.0522 | 0.0138 |
| SCHOOL | 0.0091 | 0.0109 | 0.0069 | 0.0104 | 0.0043 | 0.0083 |

| Standard Deviation | 1940 | 1950 | 1960 | 1970 | 1980 | 1990 |
|--------------------|-------|-------|-------|-------|-------|-------|
| GTS | 0.088 | 0.053 | 0.039 | 0.046 | 0.022 | 0.025 |
| ITS | 0.094 | 0.087 | 0.028 | 0.048 | 0.040 | 0.036 |
| STAR | 0.079 | 0.041 | 0.031 | 0.025 | 0.020 | 0.026 |
| SPERSTAR | 0.100 | 0.066 | 0.040 | 0.033 | 0.022 | 0.024 |
| STAR(AV) | 0.081 | 0.051 | 0.033 | 0.027 | 0.020 | 0.024 |
| SPER(AV) | 0.088 | 0.055 | 0.031 | 0.022 | 0.017 | 0.030 |
| TRANSLOG | 0.184 | 0.167 | 0.130 | 0.131 | 0.105 | 0.096 |
| LIB | 0.314 | 0.288 | 0.225 | 0.237 | 0.206 | 0.216 |
| SCHOOL | 0.113 | 0.095 | 0.066 | 0.050 | 0.035 | 0.023 |

of the variables. The second panel reports the average annual growth rate for each decade, and the growth rate for the overall sample period 1940 to 1990. We note that the average annual growth rate over the overall sample period is very close to 0.4 percent for all measures involving Divisia indexes. That is, the aggregate growth rate according to GTS, ITS, STAR, SUPERSTAR, STAR (AV), and SUPERSTAR (AV) is very similar. The annual growth rate of human capital according to the direct TRANSLOG estimation is 2 percent. The LIB measure grows at an average rate of 1.4 percent while the growth rate of the average years of SCHOOLing is 0.8 percent per year.

Hence, we note that, even though the average years of schooling of the United States labor force increased at 0.8 percent per year over the course of half a century (which implies an

Appendix Table 2. Cross-correlations (stacked variables).

| | GTS | ITS | STAR | SUPER STAR | STAR (average) | SUPER (average) | TRANS LOG | LIB | SCHOOL |
|----------|------|------|------|---------------|-------------------|--------------------|--------------|------|--------|
| HKGTS | 1.00 | 0.84 | 0.98 | 0.96 | 0.98 | 0.90 | 0.94 | 0.61 | 0.97 |
| HKITS | | 1.00 | 0.86 | 0.84 | 0.87 | 0.87 | 0.76 | 0.37 | 0.82 |
| HKSTRMIN | | | 1.00 | 0.96 | 1.00 | 0.92 | 0.92 | 0.60 | 0.96 |
| HKSSMIN | | | | 1.00 | 0.98 | 0.96 | 0.93 | 0.60 | 0.95 |
| HKSTRAVG | | | | | 1.00 | 0.93 | 0.94 | 0.61 | 0.97 |
| HKSSAVG | | | | | | 1.00 | 0.80 | 0.49 | 0.86 |
| TRANSLOG | | | | | | | 1.00 | 0.62 | 0.98 |
| HKW0 | | | | | | | | 1.00 | 0.62 |
| SCHAVL | | | | | | | | | 1.00 |

overall increase of 80 percent), human capital increased at a rate of only 0.4 percent (which implies an overall increase of 40 percent.)

The table also reports the growth rates of each of the measures decade by decade. We note that there are substantial differences in the way the various variables behave over time.

The third panel of Appendix Table 1 shows the behavior of the cross-sectional standard deviation of each of the measures over time (that is, this panel reports, for each year, the standard deviation of the 48 state estimates for each measure). Because our variables are all measured in logs, the measure of dispersion is similar to the coefficient of variation in that it is invariant to proportional changes in the levels of the variables. The important point about this panel is that the behavior varies a lot from measure to measure. If we look, for example, at the prediction of the behavior of dispersion during the 1980s, we see that SCHOOL, TRANSLOG, and ITS predict a decline. The rest of the measures predict an increase. In summary, a quick look at the behavior of the U.S. aggregate values for all the human capital measures suggests that it matters a lot which measure we use.

Appendix Table 2 reports the cross-correlation of all nine measures, when each one of them is stacked on one large column vector. The important point about this table is that the correlations are very high (above 90 percent for almost all measures). The two possible exceptions are the ITS and LIB. The correlation between ITS and the rest of the measures is below 0.9 in all cases. The correlations between LIB and the rest of the variables is much smaller and they range from 0.36 (ITS) and 0.62 (School).

Appendix Table 3 decomposes the cross-correlations by decade. The pattern of correlations is quite interesting. First, the cross-correlations involving ITS and LIB are quite low. Second, the number of correlation coefficients about 0.90 seems to fall over time. In other words, even though all measures yield a similar human capital index for 1940, the same is not true for 1980 or 1990. Hence, it matters a lot which measure we choose. The correlation between our favorite measure, GTS; and SCHOOL falls from 0.93 in 1940 to a low 0.83 in 1970 and then it goes back up to 0.85 in 1990.

Third, the correlation between the TRANSLOG variable and the rest of the measures deteriorates with time. In particular, the correlations for 1980 and 1990 are very close to zero for some variables. The correlation between TRANSLOG and SCHOOL for 1990 is also very low: 0.63.

Appendix Table 3. Cross-correlations by decade.

| 1940 | GTS | ITS | STAR | SUPER | STAR(A) | SUP(AV) | TRANSL | LIB | SCHOOL |
|---------|-----|-------|------|-------|---------|---------|--------|------|--------|
| GTS | 1 | 0.96 | 0.99 | 0.90 | 0.99 | 0.89 | 0.94 | 0.56 | 0.96 |
| ITS | | 1.00 | 0.96 | 0.85 | 0.96 | 0.85 | 0.89 | 0.48 | 0.88 |
| STAR | | | 1.00 | 0.91 | 1.00 | 0.91 | 0.95 | 0.59 | 0.96 |
| SUPER | | | | 1.00 | 0.93 | 1.00 | 0.90 | 0.61 | 0.85 |
| STAR(A) | | | | | 1.00 | 0.93 | 0.95 | 0.59 | 0.94 |
| SPU(AV) | | | | | | 1.00 | 0.89 | 0.63 | 0.85 |
| TRANSL | | | | | | | 1.00 | 0.60 | 0.94 |
| LIB | | | | | | | | 1.00 | 0.62 |
| 1950 | | | | | | | | | |
| GTS | 1 | 0.92 | 0.95 | 0.90 | 0.97 | 0.89 | 0.92 | 0.50 | 0.94 |
| ITS | | 1.00 | 0.89 | 0.86 | 0.92 | 0.85 | 0.87 | 0.57 | 0.88 |
| STAR | | | 1.00 | 0.88 | 0.97 | 0.87 | 0.84 | 0.42 | 0.87 |
| SUPER | | | | 1.00 | 0.95 | 1.00 | 0.89 | 0.42 | 0.84 |
| STAR(A) | | | | | 1.00 | 0.94 | 0.93 | 0.46 | 0.92 |
| SPU(AV) | | | | | | 1.00 | 0.87 | 0.43 | 0.82 |
| TRANSL | | | | | | | 1.00 | 0.45 | 0.96 |
| LIB | | | | | | | | 1.00 | 0.45 |
| 1960 | | | | | | | | | |
| GTS | 1 | 0.70 | 0.98 | 0.95 | 0.98 | 0.94 | 0.91 | 0.27 | 0.93 |
| ITS | | 1.00 | 0.74 | 0.67 | 0.73 | 0.72 | 0.54 | 0.37 | 0.61 |
| STAR | | | 1.00 | 0.96 | 1.00 | 0.96 | 0.90 | 0.39 | 0.93 |
| SUPER | | | | 1.00 | 0.97 | 0.99 | 0.90 | 0.35 | 0.89 |
| STAR(A) | | | | | 1.00 | 0.97 | 0.91 | 0.39 | 0.92 |
| SPU(AV) | | | | | | 1.00 | 0.85 | 0.39 | 0.86 |
| TRANSL | | | | | | | 1.00 | 0.35 | 0.96 |
| LIB | | | | | | | | 1.00 | 0.32 |
| 1970 | | | | | | | | | |
| GTS | 1 | 0.51 | 0.84 | 0.87 | 0.85 | 0.78 | 0.79 | 0.03 | 0.83 |
| ITS | | 1.00 | 0.61 | 0.57 | 0.60 | 0.61 | 0.35 | 0.13 | 0.37 |
| STAR | | | 1.00 | 0.98 | 0.99 | 0.96 | 0.77 | 0.21 | 0.87 |
| SUPER | | | | 1.00 | 0.99 | 0.92 | 0.86 | 0.14 | 0.93 |
| STAR(A) | | | | | 1.00 | 0.96 | 0.80 | 0.22 | 0.89 |
| SPU(AV) | | | | | | 1.00 | 0.61 | 0.30 | 0.80 |
| TRANSL | | | | | | | 1.00 | 0.00 | 0.90 |
| LIB | | | | | | | | 1.00 | 0.17 |
| 1980 | | | | | | | | | |
| GTS | 1 | −0.04 | 0.97 | 0.93 | 0.97 | 0.58 | 0.69 | 0.29 | 0.89 |
| ITS | | 1.00 | 0.05 | 0.12 | 0.03 | 0.20 | 0.03 | 0.24 | 0.03 |
| STAR | | | 1.00 | 0.97 | 1.00 | 0.62 | 0.72 | 0.29 | 0.93 |
| SUPER | | | | 1.00 | 0.97 | 0.73 | 0.63 | 0.33 | 0.90 |
| STAR(A) | | | | | 1.00 | 0.62 | 0.71 | 0.28 | 0.92 |
| SPU(AV) | | | | | | 1.00 | −0.04 | 0.40 | 0.40 |
| TRANSL | | | | | | | 1.00 | 0.11 | 0.87 |
| LIB | | | | | | | | 1.00 | 0.20 |

(continued)

Appendix Table 3. Continued.

| 1990 | GTS | ITS | STAR | SUPER | STAR(A | SUP(AV) | TRANSL | LIB | SCHOOL |
|---------|-----|------|------|-------|--------|---------|--------|-------|--------|
| GTS | 1 | 0.31 | 0.96 | 0.84 | 0.98 | 0.48 | 0.35 | 0.34 | 0.85 |
| ITS | | 1.00 | 0.43 | 0.39 | 0.41 | 0.20 | 0.15 | −0.02 | 0.49 |
| STAR | | | 1.00 | 0.75 | 0.99 | 0.35 | 0.46 | 0.26 | 0.93 |
| SUPER | | | | 1.00 | 0.81 | 0.85 | −0.10 | 0.38 | 0.61 |
| STAR(A | | | | | 1.00 | 0.42 | 0.42 | 0.29 | 0.92 |
| SPU(AV) | | | | | | 1.00 | −0.61 | 0.28 | 0.13 |
| TRANSL | | | | | | | 1.00 | 0.04 | 0.64 |
| LIB | | | | | | | | 1.00 | 0.26 |

Fourth, for the later years, the two measures involving SUPERSTAR orderings (SUPERSTAR and SUPERSTAR (AV)) are poorly correlated with the rest of the variables.

Finally, the optimal star configuration is very correlated with GTS for all years. The only possible exception seems to be the 1970s, whose correlation is 0.83. For the rest of the years the correlation is close or above to 0.96.

Appendix Table 4 shows the autocorrelations for each of the variables estimated over the 10-year periods for which we have data. The first thing to note is that the autocorrelation for the average years of schooling is close or above 0.9 for all decades. This is not true for any other estimate. For example, even though the autocorrelation for GTS is quite large, it is only above 0.9 in the first decade. The correlation between 1960 and 1970 is 0.88, between 1970 and 1980 it is 0.64, and in the last decade of the same period it is. The autocorrelations for all the other measures range between zero and 0.99.

The conclusion from this empirical analysis of the various measures is that the various measures proposed and described throughout the article are very different and that finding the optimal configuration is important. And this is what we attempted to do in this article.

Appendix B. Data Description and Sources

To compute human capital indexes across states of the United States, we use census data for all census years starting in 1940. We use the Public Use Microdata provided by the Census Bureau. The microsamples include information on the schooling, earnings, hours and weeks worked, and employment status of a (practically speaking) random sample of roughly one out of every 100 Americans in each of the census years 1940, 1950, 1960, 1970, 1980, and 1990.¹⁴

From the micro data, we compute the educational attainment distribution (among civilians aged 25 to 65) for each state at each of the six dates and the average weekly earnings for each schooling group at each date.

Each state's educational attainment distribution is estimated by dividing its civilian labor force aged 25 to 65 into seven schooling categories:

0. No schooling

Appendix Table 4. Autocorrelations.

A. Optimum measures

| 1940 | 1950 | 1960 | 1970 | 1980 | 1990 |
|-----------------------------------|------|------|------|-------|------|
| Geographic traveling salesman: | | | | | |
| 1.00 | 0.93 | 0.90 | 0.64 | 0.73 | 0.49 |
| | 1.00 | 0.87 | 0.67 | 0.78 | 0.62 |
| | | 1.00 | 0.74 | 0.83 | 0.59 |
| | | | 1.00 | 0.64 | 0.53 |
| | | | | 1.00 | 0.79 |
| | | | | | 1.00 |
| Intertemporal traveling salesman: | | | | | |
| 1.00 | 0.84 | 0.57 | 0.33 | −0.15 | 0.28 |
| | 1.00 | 0.63 | 0.26 | −0.03 | 0.33 |
| | | 1.00 | 0.41 | −0.10 | 0.23 |
| | | | 1.00 | 0.03 | 0.22 |
| | | | | 1.00 | 0.23 |
| | | | | | 1.00 |
| Star (minimum): | | | | | |
| 1.00 | 0.90 | 0.88 | 0.69 | 0.74 | 0.59 |
| | 1.00 | 0.88 | 0.74 | 0.72 | 0.55 |
| | | 1.00 | 0.82 | 0.86 | 0.67 |
| | | | 1.00 | 0.79 | 0.64 |
| | | | | 1.00 | 0.88 |
| | | | | | 1.00 |
| Superstar (minimum): | | | | | |
| 1.00 | 0.94 | 0.90 | 0.70 | 0.58 | 0.30 |
| | 1.00 | 0.94 | 0.74 | 0.63 | 0.33 |
| | | 1.00 | 0.86 | 0.69 | 0.31 |
| | | | 1.00 | 0.76 | 0.44 |
| | | | | 1.00 | 0.66 |
| | | | | | 1.00 |

B. Nonoptimum measures

| 1940 | 1950 | 1960 | 1970 | 1980 | 1990 |
|-----------------|------|------|------|------|------|
| Star (average): | | | | | |
| 1.00 | 0.94 | 0.90 | 0.71 | 0.74 | 0.57 |
| | 1.00 | 0.93 | 0.77 | 0.77 | 0.62 |
| | | 1.00 | 0.84 | 0.85 | 0.65 |
| | | | 1.00 | 0.80 | 0.67 |
| | | | | 1.00 | 0.86 |
| | | | | | 1.00 |

(continued)

Appendix Table 4. Continued.

B. Nonoptimum measures

| 1940 | 1950 | 1960 | 1970 | 1980 | 1990 |
|-----------------------------------|------|------|------|------|-------|
| Superstar (average): | | | | | |
| 1.00 | 0.94 | 0.86 | 0.52 | 0.08 | −0.05 |
| | 1.00 | 0.92 | 0.60 | 0.18 | −0.01 |
| | | 1.00 | 0.74 | 0.25 | −0.02 |
| | | | 1.00 | 0.51 | 0.32 |
| Translog estimates: | | | | | |
| 1.00 | 0.97 | 0.95 | 0.83 | 0.85 | 0.59 |
| | 1.00 | 0.95 | 0.83 | 0.84 | 0.59 |
| | | 1.00 | 0.91 | 0.89 | 0.65 |
| | | | 1.00 | 0.89 | 0.66 |
| | | | | 1.00 | 0.83 |
| | | | | | 1.00 |
| Labor-income-based human capital: | | | | | |
| 1.00 | 0.70 | 0.70 | 0.46 | 0.52 | 0.45 |
| | 1.00 | 0.71 | 0.41 | 0.50 | 0.40 |
| | | 1.00 | 0.71 | 0.75 | 0.63 |
| | | | 1.00 | 0.65 | 0.68 |
| | | | | 1.00 | 0.66 |
| | | | | | 1.00 |
| Average years of schooling: | | | | | |
| 1.00 | 0.98 | 0.95 | 0.84 | 0.83 | 0.67 |
| | 1.00 | 0.97 | 0.87 | 0.86 | 0.71 |
| | | 1.00 | 0.91 | 0.92 | 0.74 |
| | | | 1.00 | 0.91 | 0.76 |
| | | | | 1.00 | 0.90 |
| | | | | | 1.00 |

Appendix Table 5. Correlation between human capital estimates using ad-hoc quadratic cost function (5) and derived cost function (8).

| | |
|---------------|-------|
| 1940 | 0.991 |
| 1950 | 0.897 |
| 1960 | 0.983 |
| 1970 | 0.837 |
| 1980 | 0.963 |
| 1990 | 0.944 |
| Stacked 0.973 | |

1. 0 to 4 years of elementary school
2. 5 to 8 years of elementary school
3. 1 to 3 years of high school
4. High school graduate
5. 1 to 3 years of college
6. College graduate or more

“Years of schooling” refers to the highest grade completed.¹⁵ Nursery school and kindergarten are not counted as grades, so an individual qualifies for our no-schooling category if he attended nursery school or kindergarten or even if he attended but did not complete first grade.

We will take a worker’s average weekly earnings (annual earnings divided by weeks worked, both for the year prior to the census) to be his or her marginal productivity. Estimates of average weekly earnings are obtained from a subsample of the aged 25 to 65 civilian labor force (this subsample will be referred to as our “earnings sample”): employed civilians who worked at least 13 weeks in the year prior to the census (the year for which earnings are reported), who were not self-employed, and who worked more than 30 usual hours per week (in the year of the census or in the year prior to the census, depending on the orientation of the census question in that year). Workers were excluded from the earnings sample if they were currently attending school or if, on average, they earned less than 67 1982 dollars per week, adjusted for “real economic growth” at 2 percent per year. Aged 25 to 65 civilians who satisfy these selection criteria form our “earnings sample.”

Appendix C. Technical Discussion of Five Approximations Used to Compute the Optimal Configuration

In Section 6 we argue that five approximations are used that allow us to arrive at an estimate of the optimal configurations. The first two approximations involve specification of the priors $\Pi(\omega)$.

Consider a two Taylor expansions of the index function F —one at economy A’s endowment point and one at economy B’s:

$$\begin{aligned}
 \log F(X) = & \log F(X_A) + \sum_{i=1}^N v_{i,A} (\log x_i - \log x_{i,A}) \\
 & + \sum_{i=1}^N \sum_{j=1}^N \frac{1}{2} \frac{\partial^2 \log F}{\partial (\log x_i) \partial (\log x_j)} \bigg|_{X=X_A} \\
 & \times (\log x_i - \log x_{i,A}) (\log x_j - \log x_{j,A}) + (\text{higher-order terms})
 \end{aligned}$$

$$\begin{aligned}
\log F(X) &= \log F(X_B) + \sum_{i=1}^N v_{i,B} (\log x_i - \log x_{i,B}) \\
&\quad + \sum_{i=1}^N \sum_{j=1}^N \frac{1}{2} \frac{\partial^2 \log F}{\partial (\log x_i) \partial (\log x_j)} \Big|_{X=X_B} \\
&\quad \times (\log x_i - \log x_{i,B}) (\log x_j - \log x_{j,B}) + (\text{higher-order terms}),
\end{aligned}$$

where “higher order terms” include third and higher-order terms.

Approximation 1

For each pair of economies A and B, consider a second-order approximation to the difference between each of their second derivatives. The slope terms $\{\gamma_{i,j,k}^{A,B}\}$ are unknown, but the econometrician has priors that are distributed as joint multivariate normal with mean zero.

Approximation 1 allows for a simple representation of the second derivative terms:

$$\frac{\partial^2 \log F}{\partial (\log x_i) \partial (\log x_j)} \Big|_{X=X_A} - \frac{\partial^2 \log F}{\partial (\log x_i) \partial (\log x_j)} \Big|_{X=X_B} = \sum_{k=1}^N \gamma_{i,j,k}^{A,B} (\log x_{k,A} - \log x_{k,B}).$$

If we evaluate $F(X)$ at either X_A or X_B , take the average of the two Taylor expansions, and invoke Approximation 1, we can compute pieces of the cost function (8) for a given configuration ω :

$$\begin{aligned}
[d_{A,B}(\omega)] &= E \left(\log \frac{\hat{H}_A(\omega)}{\hat{H}_B(\omega)} - \log \frac{H_A}{H_B} \right)^2 = \\
&\sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^N \sum_{l=1}^N \sum_{m=1}^N \sum_{n=1}^N E(\gamma_{i,j,k}^{A,B} \gamma_{l,m,n}^{A,B}) \Delta(x_A, x_B, i, j, k, l, m, n) \Delta(x_A, x_B, i, j, k, l, m, n) \\
&\equiv [\log(x_{i,A}/x_{i,B})][\log(x_{j,A}/x_{j,B})][\log(x_{k,A}/x_{k,B})] \\
&\quad \times [\log(x_{l,A}/x_{l,B})][\log(x_{m,A}/x_{m,B})][\log(x_{n,A}/x_{n,B})],
\end{aligned}$$

where economies A and B are neighbors in configuration ω . This set of six sums has N^6 terms and must be computed for each pair of adjacent economies! Approximation 2 simplifies the computation, resulting in the triple sum (10).

Approximation 2

For each pair of economies, let the priors about the third-order coefficients have mean zero. Approximate the covariance matrix with a diagonal matrix and assume that the matrix is a scalar multiple of the identity matrix. The econometrician’s degree of uncertainty about

the third-order parameters is independent of the pair of economies being compared:

$$\begin{aligned}
 E[d_{A,B}(\omega)] &= \sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^N E[(\gamma_{i,j,k}^{A,B})^2] [\log(x_{i,A}/x_{i,B})]^2 \\
 &\quad \times [\log(x_{j,A}/x_{j,B})]^2 [\log(x_{k,A}/x_{k,B})]^2 = \\
 &= E[\gamma^2] \sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^N [\log(x_{i,A}/x_{i,B})]^2 \\
 &\quad \times [\log(x_{j,A}/x_{j,B})]^2 [\log(x_{k,A}/x_{k,B})]^2. \tag{10}
 \end{aligned}$$

Approximation 2 says two things. First, the econometrician knows that the true aggregator function is not quadratic (in the log space) but is not sure about the direction of deviations from quadratic. Third-order terms may be negative or positive. Second, approximation 2 rules out certain third-order interactions in the aggregator function: this is the substance of the assumption of a diagonal covariance matrix.

Approximations 1 and 2 still do not specify the computation of $E[d_{A,B}(\omega)]$ if economies A and B are not directly linked in the configuration ω . Approximation 3 fills this gap:

Approximation 3

For two economies A and B that are directly linked to a third economy C (and therefore, according to the definition of the *unilink* space, not directly linked to each other), their expected squared bilateral comparison error $d_{A,B}(\omega)$ can be approximated as the sum of $d_{A,C}(\omega)$ and $d_{B,C}(\omega)$:

$$E[d_{A,B}(\omega)] = E[d_{A,C}(\omega)] + E[d_{B,C}(\omega)].$$

Approximation 3 rules out correlations *across links* of the bilateral comparison errors.

Approximations 1, 2, and 3 fully specify the computation of the cost function $C(\omega)$. Each term of the cost function involves a squared bilateral comparison error $d_{A,B}(\omega)$, which is computed according to equation (10) for adjacent economies or according to Approximation 3 for nonadjacent economies. The minimization Problem (9) can now be computed in a straightforward way. However, because of the tremendous size of the space $\Omega(M)$, such computations are impractical. Approximations 4 and 5 are intended to facilitate the minimization procedure.

Approximation 4

Replace the minimization over the *unilink* space with a minimization of cost over the union of Ω_1 , Ω_2 , Ω_3 , and Ω_4 :

$$\begin{aligned}
 \tilde{\omega} &= \operatorname{argmin}_{\omega \in (\Omega_1 \cup \Omega_2 \cup \Omega_3 \cup \Omega_4) \subset \Omega} C(\omega) \\
 &= \operatorname{argmin} \left[\min_{\omega \in \Omega_1} C(\omega), \min_{\omega \in \Omega_2} C(\omega), \min_{\omega \in \Omega_3} C(\omega), \min_{\omega \in \Omega_4} C(\omega) \right]. \tag{11}
 \end{aligned}$$

The minimization problem (11) still presents a formidable computation problem. The difficulty is that the GTS and ITS spaces (Ω_1 and Ω_2) are still fairly large and the cost function described by equation (10) and Approximation 3 is quite complicated. Because numeral optimization over large spaces is facilitated when the cost function is easy to evaluate, Approximation 5 attempts to simplify the evaluation of the cost function of the GTS and ITS spaces:

Approximation 5

Define a pseudocost function, to be the sum of the expected squared bilateral comparison errors with the sum taken over *adjacent economies only*. Consider the minimization of the pseudocost function over Ω_1 and Ω_2 separately:

$$\begin{aligned}\tilde{\omega}' &\equiv \operatorname{argmin}_{\omega \in \Omega_1} \tilde{C}(\omega) \\ \tilde{\omega}'' &\equiv \operatorname{argmin}_{\omega \in \Omega_2} \tilde{C}(\omega).\end{aligned}$$

Effectively, the pseudocost function differs from the actual cost function in the way that it weights the different links. To see this, consider a four-economy example. When the four economies are arranged linearly as $A-B-C-D$, the true cost function weights $d_{B,C}$ four times as heavily as $d_{A,B}$ or $d_{C,D}$. This is because $d_{B,C}$ also enters into the comparison error for pairs (A, C) , (B, D) , and (A, D) , whereas $d_{A,B}$ matters only for the (A, B) comparison. The pseudocost function, on the other hand, weights each direct link equally. Also notice that the difference between pseudocost and cost depends on the space. For the GTS and ITS spaces, pseudocost and cost can be quite different. For an element of the superstar space, on the other hand, pseudocost and cost are identical (up to the constant $2(M-1)/M^2$). Pseudocost and cost are also quite similar for an element of the star space (any difference arises from intertemporal links).

Our estimate of the optimal configuration uses the pseudocost function to economize on the searches of the GTS and ITS spaces:

$$\hat{\omega} \equiv \operatorname{argmin} \left[C(\tilde{\omega}'), C(\tilde{\omega}''), \min_{\omega \in \Omega_3} C(\omega), \min_{\omega \in \Omega_4} C(\omega) \right]. \quad (12)$$

Approximation 5 implies that we use the actual cost function to estimate minimum cost configurations for the Ω_3 and Ω_4 spaces (*star* and *superstar* respectively), use the pseudocost function to estimate minimum cost configurations for the spaces Ω_1 and Ω_2 , and use the actual cost function to choose between the spaces Ω_1 , Ω_2 , Ω_3 , and Ω_4 .

Acknowledgments

Our research has benefited from discussions with Robert Barro, Gary Becker, Jess Benhabib, Fisher Black, Bill Brainard, Jordi Gali, Jim Heckman, Hugo Hopenhayn, Dale Jorgenson, Boyan Jovanovic, Per Krussell, Robert Lucas, Yai Mundlak, Kevin Murphy, Derek Neal,

Quin Pal, Chris Sims, Nancy Stokey, Merritt Tilney, Petra Todd, Alwyn Young, and participants at seminars at the Universities of Chicago, Columbia, Pompeu Fabra, Elsa Vila-Artadi NYU, Yale, IVIE, and the NBER, IMP, and CEPR. This research started when Xavier was visiting Universitat Pompeu Fabra. The authors acknowledge the financial support of the NSF.

Notes

1. The existing literature on economic growth has used two very simple aggregators: the linear sum of all workers (that is, the aggregate employment or the aggregate labor force) and the weighted sum of all workers where the weights are the years of schooling (that is, the average years of schooling.)
2. Barro and Lee (1996) also discusses the quality of schooling across countries. For a discussion of cross-sectional differences in schooling quality, see Woessmann (2000).
3. In principle, there are an endless number of ways to categorize workers: years of schooling, type of schooling, occupation, industry, sex, age, and race are some possibilities. In later sections we categorize workers from United States census state data by educational attainment and gender.
4. The reader can verify that Laspeyres and Paasche approximations are first order by showing that they are without error if and only if the true aggregator function is a Cobb-Douglas function. The Divisia approximation is said to be second order because it is without error if and only if the true aggregator function is a translog function. Of course, the Cobb-Douglas function is a first-order linear approximation of the translog function in the log space (that is, the Cobb-Douglas is a particular case of (1) when $\beta = 0$).
5. Even though 1941 and 1940 are chronologically next to each other, it is not obvious that we want to treat them this way when we compute the growth rate of human capital between these two periods. We will argue that, for the Divisia index to be a good approximation to the true aggregator, the inputs of the two economies being compared must be “similar.” It is not necessarily true the economy’s inputs in two consecutive years are more similar. Therefore, the chronological order is not necessary the correct order (although it will tend to be correct when the components of the input vector change little over time).
6. Instead of using the *superstar* method, we could construct a panel data set of human-capital numbers by computing the cross-section of numbers for a base period using the *star system* (which allows for cross-sectional comparisons for that particular year) and then compute the time series for each economy. The problem with this is that there are a lot of possible ways of doing this. For example, to compare the stock of human capital in California and Illinois in 1990, we could compute a cross-section of human capitals for 1940 and then the time series for California and Illinois, or we could compute the cross-section for 1990, which would allow for a direct comparison, or we could compute the cross-section for 1970, 1950, 1960, or 1980 and then the time series for California. We estimated a variety of them and found the results to be very unreliable in that they were very sensitive to the exact choices of base years.
7. Alternatively, we could report a 288 by 288 matrix of bilateral comparisons and let the reader choose among the exceedingly large number of possible multilateral comparisons.
8. Optimal is defined in Section 6.
9. With a few hundred economies, as in our example, searches of the unilink space could take billions of years even with a very fast computer. Even the best computer algorithms need thousands of years.
10. The key difference between the method we described here and Kravis’s star system described in Section 4 is that here we find the star that minimizes the cost function rather than arbitrarily imposing a star or base economy or by arbitrarily computing the average human capital for all possible star economies. In principle, we could allow for different geographical stars within the same year. For example, we could link all the southern states in one star and the same thing for all the midwestern, eastern and western states. We could then link each star by the two states that minimize cost. We conjecture that this strategy would deliver very good indexes in the sense that their cost would be very low. The problem is that the construction of these multiple stars is very complicated. Furthermore, there is no reason for breaking the United States into four regions only. The choice of the number of stars would complicate the problem beyond our ability to solve it.
11. For example, one might use the simple Euclidean distance metric in the log space rather than weighting category growth rates by category income shares as in equation (5).

12. We have only been able to put a lower bound on the minimum cost from the star space.
13. The existence of temporary measurement error in the human capital estimates could be the explanation for this high speed of convergence.
14. The 1940, 1950, and 1960 PUMS are 1/100. To economize on computing resources, we work with random subsamples of the later PUMS, arriving at a 1/1000 sample for 1970 and 1/200 samples for 1980 and 1990. The 1970 subsample was provided by the Census Bureau who derived it from their 5 percent state sample. Our 1980 and 1990 subsamples were constructed from the larger 5 percent sample by taking only those households whose subsample number had a ones digit equal to 2.
According to the Census Bureau, the 1960, 1970, and 1980 PUMS are self-weighting samples, and we treated them as such in our computations. We also treat our extracts of the 1940 and 1950 PUMS as self-weighting samples, noting that we extract all persons from the raw 1940 data files and only sample line persons from the raw 1950 data files.
The 1990 PUMS is clearly not a self-weighted sample, so all of our computations for that year weight by the Census Bureau's estimate of the inverse of each person's sampling probability (columns 18 to 21 of the person record). For example, the regression criteria is to minimize a weighted sum of the squared residuals, where each person's residual is weighted by the inverse of his sampling probability.
15. Until 1990, it is not clear whether a response such as "6 years of college" means that the person obtained a bachelor's degree and worked for two years for a higher degree or whether he worked six years to obtain the bachelor's degree.

References

- Barro, R. J. (1991). "Economic Growth in a Cross-Section of Countries." *Quarterly Journal of Economics* 106, 407–442.
- Barro, R. J., and J. W. Lee. (1993). "International Comparisons of Educational Attainment." *Journal of Monetary Economics* 32, 363–394.
- Barro, Robert J., and Jong-Wha Lee. (1996). "International Measures of Schooling Years and Schooling Quality." *American Economic Review* 86, 218–223.
- Barro, R. J., and X. Sala-i-Martin. (1992). "Convergence." *Journal of Political Economy* 100, 223–251.
- Diewert, W. E. (1976). "Exact and Superlative Index Numbers." *Journal of Econometrics* 4, 115–145.
- Diewert, W. E. (1987). "Index Numbers," In John Eatwell et al (eds.), *The New Palgrave: A Dictionary of Economics* (pp. 767–780). London: Macmillan.
- Fisher, I. (1922). *The Making of Index Numbers*. Boston: Houghton Mifflin.
- Garey, M. R., and D. S. Johnson. (1979). *Computers and Intractability: A Guide to the Theory of NP-Completeness*. San Francisco: Freeman.
- Kirkpatrick, S. et al. (1983). "Optimization by Simulated Annealing." *Science* 220.
- Kravis, I. B. (1984). "Comparative Studies of National Income and Prices." *Journal of Economic Literature* 22, 1–39.
- Kyriacou, George A. (1992). "A Cross-Country Estimation of an Aggregate Production Function with Human Capital." Working paper. Central Bank of Cyprus.
- Laspeyres, E. (1871). "Die Berechnung einer mittleren Warenpreissteigerung." *Jarbücher für Nationalökonomie und Statistik* 16, 296–314.
- Lucas, R. E. (1988). "On the Mechanics of Economic Development." *Journal of Monetary Economics* 22, 3–42.
- Mulligan, C., and X. Sala-i-Martin. (1997). "A Labor Income Based Measure of Human Capital." *Japan and the World Economy* 9, 159–191.
- Paasche, H. (1874). "Über die Preisentwicklung der letzten Jahr nach den Hamburger Börsennotirungen." *Jarbücher für Nationalökonomie und Statistik* 23, 168–178.
- Press, William H., et al. (1988). *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge: Cambridge University Press.
- Summers, R., and A. Heston. (1991). "The Penn World Table, Mark 5: An Expanded Set of International Comparisons, 1950–1988." *Quarterly Journal of Economics* 106, 327–368.
- Törnquist, L. (1936). "The Bank of Finland's Consumption Price Index." *Bank of Finland Monthly Bulletin* 10, 1–8.
- Woessmann, L. (2000). "Quality of Schooling and Human Capital." Mimeo, Kiel Institute of World Economics.