LONGITUDINAL EMPLOYER - HOUSEHOLD DYNAMICS

INFORMATIONAL DOCUMENT NO. ID-2003-04

A Layman's Guide to the LEHD Human Capital Measures

Date	:	January 2003
Prepared by	:	LEHD
Contact	:	U.S. Census Bureau, LEHD Program FB 2138-3 4700 Silver Hill Rd. Suitland, MD 20233 USA

A Layman's Guide to the LEHD Human Capital Measures

What they are

The human capital measures can be thought of as the market value of the portable component of an individual's skill. They have two components: an individual or person effect, which does not vary over time (θ) and a component based on labor market experience $(x\beta)$. The individual effect includes some factors that are often observable to the statistician, such as years of education and sex; and some factors that are often not, such as innate ability, "people skills," "problem solving skills," perseverance, family background, and educational quality. The experience component[1] is directly calculated from the data, and, as such, is left censored at the start of the data period. This left-censoring is ameliorated by estimating the number of years of labor force experience an individual accumulates prior to the first appearance in the data.

Table 1: Sources of Industry Earnings Differentials

SIC	*Name*	*Industry Wage Premium*	*Premium attributable to workforce human capital*	*Premium attributable to firm wage setting policy*
	Highest Paying Industries			
62	Security, commodity, brokers and services	82%	34%	37%
67	Holding and other investments	70%	34%	27%
48	Communication	63%	7%	52%
49	Electric, gas and sanitary services	54%	0%	55%
81	Legal services	54%	18%	31%
	Lowest Paying Industries			
58	Eating and drinking places	-45%	-12%	-38%
01	Agriculture-crops	-35%	-10%	-31%
72	Personal services	-33%	-12%	-24%
79	Amusement and recreation services	-32%	-8%	-28%
70	Hotel and lodging services	-32%	-17%	-19%
54	Food stores	-30%	1%	-30%

Two points are important to make here. First, the human capital measure is not the same as a wage measure. There are three distinct components of wages: human capital, a firm effect and an unexplained residual. Because the human capital measure and the firm effect are virtually uncorrelated, when measured at the level of an individual job, an individual's earnings may be due to who they are or where they work. As we point out in our proposal, this is illustrated in Table 1. Clearly the highest paying industry – security, commodity and brokers and services – is high-paying both because it has high quality workers, and because firms within that industry pay a premium to those workers. However another highly paid industry – electricity, gas and sanitary services – has high wages entirely because firms in the industry pay its workers much higher than average. The workers themselves are of the same quality as the rest of the workforce. Similar results are evidenced when low-wage industries are analyzed in the second set of panels. Eating and drinking establishments, for example, hire workers of lower than average

[1] The notation is in matrix form: the actual matrix includes experience as a quartic.

quality and pay them less. However, another very low wage industry - food stores – actually hire workers of above average quality, but just pay them less

The second point is just how important these new measures are. Traditional surveys of workers that measure the "kitchen sink" of demographic characteristics - such as education, occupation, age, sex, marital status and even include some firm characteristics such as firm size and industry – are typically able to explain some 30% of earnings variation. Longitudinal data on workers and firms explain closer to 90% of earnings variation.

What's the intuition?

The intuitive explanation for this quantitative measure is that it captures the average market value that employers assign an individual as that individual moves from firm to firm. Note that this measure is not an abstract, disembodied skill measure – like years of education, or occupation. If, for example, an individual is a highly "skilled" blacksmith, and the market does not value this skill, the new human capital measure will be correspondingly low. If the individual is physically extremely strong, and this is of decreasing value in the marketplace, the individual will have a relatively low human capital measure. However, if, for example, the individual scores high on problem solving skills, and this is valued in the market place, then s/he will have a high human capital value.

Why they are so useful

There are two main reasons why these measures are so useful. First, we would argue that they are "better" measures of human capital in a more complex economy. The case study evidence (well illustrated by Shaw, in Box 1) suggests that years of education are simply not adequate measures of human capital in a service economy.

Second, unlike the CPS-based measures discussed in the next section, they can be calculated for all individuals in an economy – and hence for all workers in a firm. The ability to create a wide variety of measures of the human capital distribution within a firm (the mean, the median, the proportion above or below an economy-wide threshold, the range, the interaction of different types of workers) is simply unprecedented. One example of where this is powerfully used is in the NBER Conference on Research in Income and Wealth held at the Federal Reserve Board April 25/26 of this year. Briefly, the measurement community is increasingly concerned that the dominant part of productivity and growth appears to be accounted for by intangible (and unobservable) assets – and held a conference, opened by Alan Greenspan, to address this issue. Our ability to measure human capital at the basic building block of economic activity – the firm – is what led Dale Jorgenson to exclaim that we had moved the analysis of productivity to the next level.

Our experience in the steel industry, particularly our plant visits, strongly suggest that these changes have resulted in a greater demand for highly skilled labor, particularly among production workers. The primary change in the workplace is that steel mills now require much greater amounts of problem solving by their workforce, for two reasons. First, information technologies put much greater amounts of data in the hands of production workers. The production line is run from computerized pulpits, where workers monitor the line and make changes as needed. Because workers are not really physically running the line—it is done more with computerized feedback loops—they are problem solvers. In the past, foreman were the problem solvers—today much of this responsibility has shifted to production workers. Furthermore, the HRM practices have changed to give workers the power to make decisions—both as they run the line, and in off-line problem-solving meetings. As a result, production workers must have good mechanical, math, and reading skills, but in addition have problem-solving skills and communications skills.

There is no existing data set, prior to that of Abowd/Haltiwanger/Lane (AHL) that can assess whether these impressions of increased skill demand are borne out in fact across the industry. Average education levels have risen modestly over time in steel. But measuring the education level is a very inadequate measure of the improved human capital that we describe above.

In summary, the AHL study concludes that there has been substantial upgrading, and that it has occurred in both new and old facilities. In assessing these results, I reach two conclusions. First, no other data set could have assessed this question of changing skill demand– there is not sufficient steel data elsewhere and the typical measures of human capital (like education or experience) are inadequate. And these results have important policy implications. They suggest that our educational and training systems must continually meet the demand for improved education and skills. And furthermore, the skills that are demanded are often subtle and general in nature—such as problem-solving skills. Second, when combined with background information from the steel industry, I reach the conclusion that the methodology used by the AHL study to measure human capital appears to produce very sensible results. All our plant visits and data from the steel industry suggest that the skill upgrading that AHL find is real.

Source: Kathryn Shaw's comments on the Abowd/Haltiwanger/Lane final submission to the Sloan Foundation, March 1, 2002

How they are estimated

There are two key components to the estimation of these measures. The first is that it is necessary to have longitudinal data on firms and employees.[2] The second is that it is necessary to have appropriate econometric techniques.

We discuss the first in our Sloan report of March 1. The technical derivation of the econometric techniques is available in Abowd/Creecy/Kramarz (ACK), but was initially developed by Abowd/Kramarz/Margolis (AKM, *Econometrica*, 1999). We simply explain the intuition here. The approach is similar to the methods used in breeding and statistical genetics. It is important in animal breeding to determine the contribution of, say, the bull to the milk yield of his offspring, just as it is important to determine the contribution of the cow. This could not be determined if the bull only mated with one cow. However, bulls mate with many cows, and cows mate with many bulls. The separation of the "bull" effect from the "cow" effect in determining milk yield is a problem that uses very similar techniques to the ones used by ACK.

How they compare to other measures

The main approach that has been used to estimate human capital was developed by Jorgenson, Gallop and Fraumeni (JGF). Briefly, the JGF approach incorporates data from the Censuses of Population, the Current Population Survey (CPS), and the National Income and Product Accounts (NIPA) and bases labor quality indices on cell based totals of labor inputs classified by sex, age, educational attainment, employment class, and industry. We summarize the results of two different types of comparison here.

The first "direct" approach compares the JGF indices to sectoral labor quality derived from industry averages of our human capital measure for the period 1995-1998. JGF formally define labor quality as the ratio of the total volume of labor to hours worked, where volume is measured by a constant quality index of labor quantity. The LEHD measure of industry average human capital follows essentially the same logic, where the measure of labor volume is also based on a constant quality human capital measure, and where total employment substitutes for total hours worked. Neither approach is completely satisfactory, because while LEHD data cannot measure hours worked, the JGF constant quality index of labor quality confounds firm heterogeneity with person heterogeneity.

We compare the growth rates in the human capital indices over the 1995-98 period using the LEHD-based and JGF approach. The within-industry growth rates are highly correlated -- the employment-weighted average of the sectoral correlations is 0.79. However, there is much higher average growth for any given industry and more cross industry variation in those growth rates in the LEHD measures compared to the JGF measures (the average growth rate for the LEHD measure over the 4 years is 0.04 with the cross industry standard deviation of 0.067 while the corresponding growth for the JGF is 0.014 with a cross industry standard deviation of 0.001).

[2] Universe data permit much more accurate measurement of the underlying distribution, and this is an advantage in the LEHD Program data; however, the techniques can be applied to properly constructed samples as long as the underlying data are longitudinal in both the employer and employee dimensions.

In what follows, we exploit cross sectional variation (across firms) in their human capital while the JGF procedure focuses on generating growth rates of human capital by industry. As such, the JGF measures are not well suited to examining within year, cross industry variation. Thus, as a second "indirect" approach we approximate the JGF labor quality indices by indices derived from predicted industry average wages obtained by regressing wages on age, education, and sex using the CPS. For this purpose, we use the same cells used by JGF. We show that the time series growth rates of these indirect measures are highly correlated with the actual JGF measures (the employment-weighted average correlation is 0.73). Thus, the CPS-based approach does a reasonable job of approximating the more sophisticated JGF measures.

Using these CPS-based measures, we compare the cross industry variation with those based on the LEHD measures for the year 1998. They are in principle comparable because both rely on regression approaches that attempt to isolate the component of wages due to individual characteristics. However, because LEHD data permit the distinction of individual from firm contributions to wages, one might not expect them to yield identical results. Workers sort non-randomly into firms based on their own characteristics--both observable and unobservable--and the characteristics of firms. Furthermore, firm wage premia--firm effects in the LEHD wage regression--are not distributed uniformly across industries. These two facts mean that there exists a strong, positive correlation between person and firm heterogeneity at the industry level (AKM) - a correlation that the JGF cell-based approach does not disentangle.

We plot the industry level aggregates for the CPS-based approach against the industry-level aggregates for the most inclusive measure of skill from the LEHD approach and report them in Figure 1. Although the levels are normalized differently, there is clearly a great deal of correlation between the two measures – indeed, the correlation is 0.76. However, there is somewhat more cross industry variation in the LEHD based measures than the CPS-based measure (the standard deviation of the former is 0.15 and the standard deviation of the latter is 0.13).

In summary, the LEHD-based measures by industry are closely related to those derived by JGF or a simpler but closely related CPS-Based procedure. However, LEHD-based measures imply greater average growth and more cross-sectional variation (in both growth rates across industries and in levels of human capital across industries within a year).

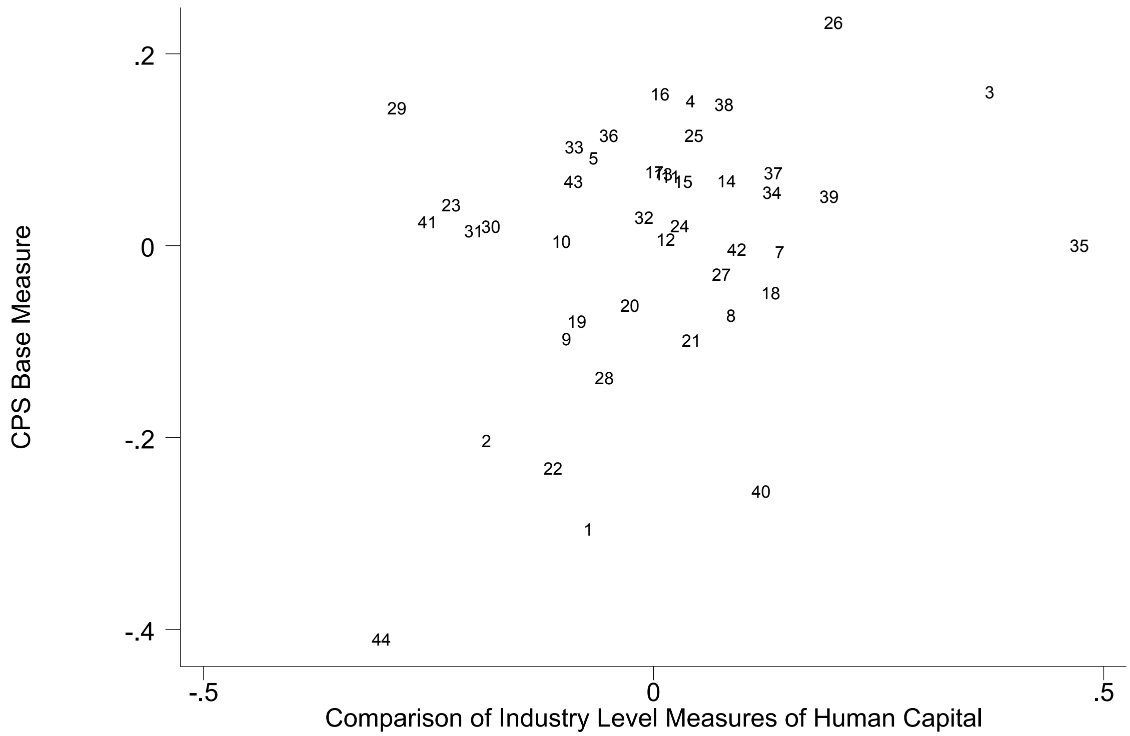
CPS Base Measure
.2
0
-.2
-.4
-.5
0
.5
Comparison of Industry Level Measures of Human Capital
26
3
16
4
38
29
36
25
33
5
37
43
15
14
34
39
23
32
41
31
30
24
12
10
35
42
7
27
18
20
8
19
9
21
28
2
22
40
1
44

Figure 1