Zurlinden, Mathias; Bolli, Thomas

Working Paper
Measurement of labor quality growth: Caused by unobservable characteristics

KOF Working Papers, No. 203

Provided in Cooperation with:
KOF Swiss Economic Institute, ETH Zurich

Suggested Citation: Zurlinden, Mathias; Bolli, Thomas (2008) : Measurement of labor quality growth: Caused by unobservable characteristics, KOF Working Papers, No. 203, ETH Zurich, KOF Swiss Economic Institute, Zurich, http://dx.doi.org/10.3929/ethz-a-005666794

This Version is available at:
http://hdl.handle.net/10419/50427

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.
You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.
If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
KOF Working Papers

Measurement of Labor Quality Growth by Unobservable Characteristics

Thomas Bolli and Mathias Zurlinden
Measurement of labor quality growth caused by unobservable characteristics*

Thomas Bolli\textsuperscript{a} and Mathias Zurlinden\textsuperscript{b}

July 2008

Abstract

The standard economy-wide indices of labor quality (or human capital) largely ignore the role of unobservable worker characteristics. In this paper, we develop a methodology for identifying the contributions of both observable and unobservable worker characteristics in the presence of the incidental parameter problem. Based on data for Switzerland over the period 1991-2006, we find that a large part of growth in labor quality is caused by shifts in the distribution of unobservable characteristics. The contributions to growth attributed to education and age are corrected downwards, if unobservable worker characteristics are taken into account. Yet the standard indices of labor quality appear to be robust to this extension.

\textit{JEL Classification:} J24, J31
\textit{Key words:} human capital, labor quality

*The authors would like to thank Elisabetta Capezzali, Marius Ley, Yves Longchamp, Barbara Rudolf, Eveline Ruoss and the participants of workshops at the ETH Zurich KOF and the "Knowledge for Growth" conference for helpful comments and discussions, and the Federal Statistical Office for providing the data. The views expressed in this paper are those of the authors and do not necessarily reflect positions of the Swiss National Bank or the ETH Zurich.

\textsuperscript{a} ETH Zurich, KOF Swiss Economic Institute, CH-8092 Zurich; Email: Bolli@kof.ethz.ch; Tel:+41 44 632 37 17
\textsuperscript{b} Economic Analysis, Swiss National Bank, CH-8022 Zurich; Email: mathias.zurlinden@snb.ch; Tel: +41 44 631 31 11; Fax: + 41 1 631 39 01.
1 Introduction

Macroeconomists have long been interested in economy-wide indices of labor quality (or human capital). The usual context is growth accounting; that is, the decomposition of output growth into the contributions of labor, capital and multi-factor productivity. Measures of labor input typically are derived from hours of workers with different education, age, gender characteristics, with wage rates serving as weights to account for differences in marginal products. The index of labor quality then is the ratio between the indices of labor input and hours worked. This standard approach is described in Jorgenson, Gollop, and Fraumeni (1987) and Bureau of Labor Statistics (1993).

Although the observable characteristics (education, age, gender) explain only a small proportion of the total variation in wages, the unobservable characteristics get little attention in the standard approach to calculating indices of labor quality. A notable exception is Abowd, Lengerman and McKinney (2002) who calculate the distribution of unobserved characteristics for the period 1992 to 1997 in U.S. data. They succeed in explaining a very large portion of the total variation in wages and attribute substantial variation to individual and employer heterogeneity.

In this paper, we add to this literature by examining the contribution of shifts in the unobserved characteristics of workers to the index of labor quality in Switzerland. The unbalanced panel data set covers the years 1991 to 2006. While the presence of the incidental parameter problem prevents us

---

\footnote{More recent studies are Aaronson and Sullivan (2001) for the U.S.; Schwerdt and Turunen (2007) for the euro area; Bell, Burriel-Llombart (2005), and Jones for the U.K.; and Bolli and Zurlinden (2008) for Switzerland.}
from estimating the individual heterogeneity consistently, we can estimate it for the average individual effect of a worker group. Based on these results, we calculate an index of labor quality that accounts for shifts in the distribution of observed and unobserved characteristics. We examine whether the standard indices of labor quality are robust to these extensions. Moreover, we compute the first-order partial indices proposed by Jorgenson et al. (1987) and examine whether the standard indices identify the sources of growth in labor quality correctly.

The paper is organized as follows. Section 2 presents the methodology. The data are described in Section 3. Sections 4 and 5 present the results and examine robustness issues. Section 6 concludes.

2 Methodology

This section first develops the methodology for calculating the index of labor quality, where shifts in the distribution of unobserved characteristics are taken into account. We then describe how the contribution of these shifts to growth in labor quality can be identified.

2.1 Calculating the index of labor quality

The methodology for calculating the index of labor quality is based on the assumption that the productivity of individual workers is reflected in their wage rates. Following the Bureau of Labor Statistics (1993), the calculation can be separated into two steps. First, earnings equations à la Mincer (1974) are estimated, and predicted wages are calculated for each individual based
on these estimates. Second, individual labor qualities are aggregated based on the methodology proposed by Jorgenson et al. (1987).

We assume that the data generating process for the natural logarithm of the real hourly wage rate \( q \) is given by

\[
\ln q_{i,t} = X_{i,t}' \beta + \alpha_i + \delta_t + \varepsilon_{i,t},
\]

where \( i \) refers to the individual and \( t \) refers to time; \( \alpha_i \) and \( \delta_t \) represent vectors of binary variables that capture unobservable heterogeneity in the dimensions individual and time; and \( X_{i,t} \) is a vector consisting of dummy variables for worker characteristics and a constant. Given the large number of individuals, estimating (1) would cause an enormous loss in degrees of freedom and would aggravate multicollinearity problems among the regressors (Baltagi, 2001). Therefore, we use the “within” estimator:

\[
\ln q_{i,t} - \ln \bar{q}_i = (X_{i,t} - \overline{X}_i)' \beta + (\delta_t - \overline{\delta}) + (\varepsilon_{i,t} - \overline{\varepsilon}_i),
\]

where \( \ln \bar{q}_i = \frac{1}{T_i} \sum_t \ln q_{i,t} \) denotes the average labor quality of individual \( i \). The averages \( \overline{X}_i, \overline{\delta} \) and \( \overline{\varepsilon}_i \) are defined analogously. Since the data set is an unbalanced panel, the number of observations per individual, \( T_i \), is varying. The “within” estimator produces consistent estimates regardless of potential correlation between explanatory variables and unobserved individual effects.\(^2\)

Following the Bureau of Labor Statistics (1993), Aaronson and Sullivan

\(^2\)The Hausman test rejects the null hypothesis that the individual effects are uncorrelated with the other explanatory variables in the model. This holds for all ten panel equations described in the text.
(2001), and Schwerdt and Turunen (2007), we estimate (2) for men and women separately to account for differences in the returns to characteristics. Furthermore we estimate (2) separately for the various education classes because the education attainment does not change after age 25 for most individuals. This gives a total of ten panel equations (2). With the gender and education characteristics dealt with in this way, $X_{i,t}$ consists of a constant and dummy variables for groups of age, where age is used as a proxy for work experience.

Given the estimated parameters $\hat{\beta}$ and $\hat{\delta}_t$, it is possible to recover the individual intercepts $\hat{\alpha}_i$:

$$\hat{\alpha}_i = \ln q_i - X_i \hat{\beta} - \hat{\delta}.$$  \hspace{1cm} (3)

These estimators are consistent if the number of observations per individual, $T_i$, approaches infinity. Since this condition is not met in our data set the presence of the incidental parameter problem prevents us from estimating the individual intercepts consistently.\footnote{For a detailed discussion of the incidental parameter problem, see e.g. Neymann and Scott (1948) and Lancaster (2000).} However, while consistency is not given, the parameter estimators are unbiased, implying that $E[\hat{\alpha}_i] = \alpha_i$ (Hsiao, 2003). Consequently, we have

$$\hat{\alpha}_i = \alpha_i + \mu_i,$$  \hspace{1cm} (4)

where $\mu_i$ is independently distributed with mean zero. Given that the number of observations per worker group $j$ can be assumed to approach infinity,
it is possible to estimate the worker group specific intercept unbiased and consistent:

\[
\lim_{N_{j,t} \to \infty} \alpha_{j,t} = \lim_{N_{j,t} \to \infty} \frac{1}{N_{j,t}} \sum_{i=1}^{N_{t}} \alpha_{i} = \lim_{N_{j,t} \to \infty} \frac{1}{N_{j,t}} \sum_{i=1}^{N_{t}} \mu_{i} = \lim_{N_{j,t} \to \infty} \hat{\alpha}_{j,t}. \tag{5}
\]

Since \( \lim_{N_{j,t} \to \infty} \frac{1}{N_{j,t}} \sum_{i=1}^{N_{t}} \mu_{i} = 0 \), it is possible to calculate predicted wage rates as

\[
\hat{q}_{j,t} = \exp(\hat{\alpha}_{j} + X_{j,t} \hat{\beta} + \hat{\delta}_{t}). \tag{6}
\]

Next, the predicted wages are used to weight the hours worked. The aggregation follows Jorgenson et al. (1987). Assuming a standard translog aggregator function, the growth rate of the quality-adjusted labor input can be calculated as

\[
\triangle \ln L_{t} = \ln \frac{L_{t}}{L_{t-1}} = \sum_{j} \left( \frac{s_{j,t} + s_{j,t-1}}{2} \ln \frac{h_{j,t}}{h_{j,t-1}} \right), \tag{7}
\]

where \( h_{j,t} \) denotes the number of total hours worked by group \( j \), and \( s_{j,t} \) is the share of labor compensation of group \( j \) in time \( t \). Finally, the growth rate of labor quality is computed as

\[
\triangle \ln Q_{t} = \triangle \ln L_{t} - \triangle \ln H_{t}, \tag{8}
\]

where \( H_{t} \) are total hours worked in the economy.
2.2 Identifying the contribution of shifts in the distribution of unobserved characteristics

To examine the effect of shifts in the distribution of unobservable characteristics, we can recalculate the index of labor quality based on predicted wage rates which do not include the contribution from the average of the unobserved characteristics, $\hat{\alpha}_j$. Thus, we have

$$\hat{q}_{j,t} = \exp(X_{j,t} \hat{\beta} + \hat{\delta}_t).$$

(9)

The modified index is calculated based on (2) and (7) to (9). In what follows, this modified index is labelled identification index while the index derived in Section 2.1 is labelled benchmark index. The difference between the benchmark index and the identification index provides a measure of the contribution of shifts in the distribution of unobserved characteristics to the index of labor quality.\(^4\)

Based on the same framework, we can decompose the index of labor quality into the partial indices for education, age and gender (and their combinations). As described by Jorgenson et al. (1987), the first-order partial indices capture the substitution between the categories of one characteristic. The indices are calculated like the total index, except that the worker groups $j$ are formed by only one characteristic instead of three.

Notice that the partial indices for education, age and gender will be biased, if they are calculated based on the model with (6), instead of (9).

\(^4\)The wages in (9) do not include unobserved characteristics. They are neither accounted for explicitly as in the benchmark methodology, nor are they included implicitly since the coefficients obtained from (2) are unbiased.
This reflects the fact that the contribution of shifts in the distribution of unobserved characteristics is captured by the partial indices of the three observable characteristics in this case. The partial indices will be more affected the stronger the correlation between the observed and unobserved characteristics.

3 Data

The data are taken from two sources: the Swiss Labor Force Survey and the Work Volume Statistic. The Federal Statistical Office (FSO) kindly provided us the micro data from these two statistics. The two statistics can be characterized as follows:

- The Swiss Labour Force Survey (SLFS) is a household survey conducted every year between April and June since 1991. The survey is representative for the permanent resident population aged 15 and older. It is based on a sampling of 33,000 households (16,000 before 2001) where each randomly selected household is interviewed over the phone five years in a row (for more information, see FSO, 2007a).

- The Work Volume Statistic (WV) is compiled from the SLFS and other sources. Data are annual and available since 1991. The WV provides more accurate data on effective working hours than the SLFS because absences due to reduced work schedules, strikes or lock-outs are taken into account (for more information, see FSO, 2007b).
Real wage rates are computed by deflating nominal hourly wages with the consumer price index. Nominal wage rates, in turn, are computed by dividing nominal earnings by hours worked. Observations of real hourly wage rates above 100 CHF are excluded from the sample because they seem to be more prone to measurement errors. Missing values are replaced by the average value of the group.

In the benchmark calculations of labor quality, three worker characteristics are considered: education, age and gender. There are five categories of education ("minimal school level", "apprentice and vocational school", "university entrance certificate", "higher vocational training", "university degree"), five age groups ("15-24", "25-39", "40-54", "55-64", "65 and more") and the two genders ("male", "female"). For some calculations, the number of categories is expanded (see Section 5).

4 Results

Based on equations (2), (3) and (6) to (8), and the data described in Section 3, we can calculate the labor quality index which accounts for changes in the distributions of observed characteristics (education, age, gender) and unobserved characteristics. Figure 1 shows this index ("Benchmark") from 1991 to 2006. The index grows by 7.1% over these 15 years, which corresponds to an average growth rate of 0.46% per year. Splitting up the sample reveals that growth is highest in the early 1990s, slows down in the second half of the decade, and speeds up again after the year 2000. The average growth rates for the sub-samples are 0.62% between 1991 and 1995, 0.26% between
1995 and 2000 and 0.52% between 2000 and 2006.

Figure 1: Indices of labor quality for various methodologies

Figure 1 also shows the indices calculated based on the methodologies proposed by Jorgenson et al. (1987) and the Bureau of Labor Statistics (1993). Jorgenson et al. (1987) use the average real wage of a worker group as a measure for labor quality. The Bureau of Labor Statistics (1993) proposes to estimate Mincerian wage equations. In contrast to (1), the presence of unobserved individual heterogeneity is not taken into account (i.e. $\alpha_i = \alpha$). Comparing these standard indices to our benchmark index reveals three points:

First, the adjustment for shifts in unobserved characteristics affects growth in labor quality. The benchmark index grows more rapidly than the index based on the method by the Bureau of Labor Statistics, and less rapidly than the index based on the method by Jorgenson et al. This implies that
the standard methodologies capture the shifts in unobserved characteristics imperfectly.

![Figure 2: Partial indices of labor quality](image)

Second, the size of the correction is moderate, suggesting that the standard indices are quite robust to the adjustment for shifts in the distribution of unobserved heterogeneity.

Third, the correction is more pronounced in the case of the method by the Bureau of Labor Statistics than in that of the method by Jorgenson et al.

It is noteworthy that the methodologies by Jorgenson et al. and the Bureau of Labor Statistics both account for some effects of the shifts in the distribution of unobserved characteristics. As described above, Jorgenson et al. use the average wage rate of a worker group as a measure of labor qual-
ity. These averages reflect both observed and unobserved characteristics and therefore the resulting index is likely to capture a substantial portion of the shifts in the distribution of unobserved characteristics. The method by the Bureau of Labor Statistics, in turn, is based on estimates of Mincerian wage equations, where the presence of unobserved individual heterogeneity is not taken into account. Consequently, the coefficients can be expected to pick up some of the effects of the omitted variables, depending on the strength of the correlation between observed and unobserved characteristics.

Turning to the first-order partial indices depicted in Figure 2, we note that the partial index of education grows by 6.3% between 1991 and 2006, implying that the substitution between education classes capture 0.41pp of labor quality growth each year. This is most of the average 0.46% per year. The second largest contribution is captured by the substitution between age classes which adds 0.19pp per year. The substitution between men and women is negligible (-0.04pp per year).\(^5\)

There are two possible explanations for the robustness of traditional labor quality indices to the adjustment for shifts in the distribution of unobserved characteristics. Either the impact of these shifts is not large, or the substitution between the worker classes considered captures the effect of these shifts reasonably well. To assess which of these two explanations is valid, we calculate the identification index described in Section 2.2. Figure 3 shows the identification index together with the benchmark index. The identification index grows by 4.7% from 1991 to 1996, corresponding to an average growth

\(^5\)The first-order partial indices of education, age and gender do not add up to the benchmark index because the second-order and third-order effects are not considered.
rate of 0.31% per year. The difference between the two series displayed in Figure 3 is substantial and suggests that labor quality growth caused by shifts in the distribution of unobserved characteristics is economically significant. Abowd et al. (2002) find too that the main driver of labor quality growth in the U.S. between 1992 and 1997 have been shifts in the distribution of unobserved characteristics.

To examine the implication of our results for the first-order partial indices of the observable characteristics, Figure 4 shows the decomposition of both the identification index and the benchmark index. We can see that the labor quality growth captured by substitution between the classes of education is lower if unobserved heterogeneity is held constant. The difference is 0.9pp over the full period. The partial indices for age suggest that the impact of

**Figure 3:** Indices of labor quality for the benchmark and the identification methodology
the substitution between age classes is overestimated as well. The size of the correction is 1.7pp. Figure 4 further shows that the labor quality growth caused by the substitution between men and women is identical in both cases.

5 Robustness

This section examines the robustness of our benchmark results with respect to alternative assumptions. The results are presented in graphs. The benchmark series are given for comparison.
5.1 Inclusion of additional worker characteristics

The benchmark index assumes that allowing for substitution between the worker groups formed by education, age and gender is sufficient to capture all changes in labor quality. In order to test this assumption, we use two additional characteristics to form worker groups: the economic sector and the employment status. We consider three different sectors (“primary”, “secondary”, “tertiary”) and two forms of the employment status (“full time”, “part time”). To prevent the number of worker per group from falling too low, the effects of these additional characteristics are examined separately.

Figure 5: Index of labor quality: set of worker characteristics expanded

Figure 5 shows that our index is affected by the inclusion of the additional characteristics. The average growth rates of the two alternative indices (“sectors expanded”, “part-time expanded”) are slightly lower than those of the
It is interesting to compare these effects to those that result if the methodology of the Bureau of Labor Statistics (1993) is applied. The inclusion of the economic sector and the employment status has qualitatively the same impact independent of the methodology. The quantitative difference is substantial, however. The inclusion of economic sectors reduces the benchmark index by 0.2% and the index based on the Bureau of Labor statistics approach by 0.4%. The correction is about two thirds of the size for the inclusion of the part-time dummy as well. This suggests that the benchmark index is more robust to the inclusion of additional variables than the index based on the method by the Bureau of Labor Statistics. The reason is that under the method by the Bureau of Labor Statistics some of the shifts in unobserved heterogeneity are captured by the additional variable. Since the methodology underlying the benchmark index already includes this effect, the correction is smaller.

5.2 Definition of the workforce

Our benchmark calculations are based on the stock of employed persons. We have excluded self-employed, apprentices and family-workers from our sample, because the assumption that the wage rate reflects the marginal product of labor is questionable for these groups. The results of calculating the benchmark index for all workers (including self-employed, apprentices and family-workers) is shown in Figure 6. The difference accumulated over 15 years amounts to 0.5%. This is a modest difference, and therefore we
conclude that the benchmark index is a valid proxy for the development of the index of labor quality of the working population in Switzerland.

5.3 Firm-specific effects

Abowd et al. (2002) argued that the organizational structure and the management skills of a firm cause differences in productivity and wages. Since these effects are not caused by the quality of labor, equation (1) should be estimated including a firm-specific intercept. Because our data set does not provide information on firm heterogeneity, this cannot be done. In order to test the robustness of our findings to this inaccuracy, we reestimate (1) with dummy variables for the twelve economic sectors as instruments. The index of labor quality based on these estimates is shown in Figure 6. The
difference to the benchmark index is small, implying that the correction for 
firm-specific effects does not have a substantial impact on our index. This is 
in line with the finding of Abowd et al. (2002), who show that most of the 
wage differences are caused by individual heterogeneity.

6 Conclusion

In this paper, we have presented a methodology that enables us to calculate 
the growth of labor quality if shifts in the distribution of unobserved char-
acteristics are accounted for. We find that the average growth rate of labor 
quality in Switzerland between 1991 and 2006 is 0.46pp. This is similar to 
the rates that result from applying the standard methodologies proposed by 
Jorgenson et al. (1987) and the Bureau of Labor Statistics (1993). This im-
plies that policy implications based on the standard indices are valid, even 
though the methodology entails a bias.

The results differ in respect to the sources of growth though. We show 
that a large part of the growth in labor quality can be attributed to shifts 
in the distribution of unobserved characteristics. Consequently the impact of 
changes in education and age is diminished substantially. The contribution 
of gender is not affected. This implies that the interpretation of unadjusted 
partial indices is questionable.
References


Abowd, John M., Francis Kramarz, and David N. Margolis (1999), ”High wage workers and high wage firms,” *Econometrica* 67(2), 251–333.


Neymann, Jerzy, and Elizabeth Scott (1948), ”Consistent estimates based on partially consistent observations,” *Econometrica* 16(1), 1–32.


### A Appendix: Tables
Table 1: Number of observations used for the estimation of Mincerian equations

<table>
<thead>
<tr>
<th>Educational Level</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimal School Level</td>
<td>16,145</td>
<td>20,840</td>
</tr>
<tr>
<td>Apprentice and Vocational School</td>
<td>54,844</td>
<td>57,240</td>
</tr>
<tr>
<td>University Entrance Certificate</td>
<td>6,596</td>
<td>10,876</td>
</tr>
<tr>
<td>University Degree</td>
<td>20,955</td>
<td>9,177</td>
</tr>
<tr>
<td>Higher Vocational School</td>
<td>15,249</td>
<td>9,674</td>
</tr>
</tbody>
</table>