QUALITY-ADJUSTED LABOUR PRODUCTIVITY MEAURES: AN APPLICATION TO NEW ZEALAND

Thomas McNaughton

*National Accounts,
Statistics New Zealand, Wellington*

Abstract

In March 2006, Statistics New Zealand released its first official productivity series. As part of the ongoing enhancements to the series, the potential for quality-adjusting the Labour Input Index has been examined. A quality-adjusted series is generally considered to provide the most representative measure of productivity. Put simply, the process acknowledges that workers are not homogenous, and as such, have different skill levels. Not only does this provide a more accurate measure of labour input, but it can also provide insight into the effects that changes in labour composition have on productivity. In practice, this process is undertaken by cross-classifying labour according to various characteristics, such as educational attainment and training, with the relative productivity levels of different groups being reflected through a wage-based weighting process. This paper evaluates the various theories behind quality-adjustment, the success of its international applications, and the potential for an internationally comparable adjustment to be introduced into the current productivity serie.

Introduction

Productivity is an important measure of the efficiency of an economy. It is calculated as the ratio of output to one or more inputs, representing the growth in output not accounted for by the growth of an input or inputs.

In March 2006, Statistics New Zealand released its first productivity series. This series, running from 1988-2005, uses capital and labour as inputs, and value added as the output measure.[1] While this series was a very significant release, there are a number of additional steps that are needed to improve the accuracy and international comparability of the series.[2]

One of these initiatives is to quality-adjust the Labour Input Index (LII). A quality-adjusted LII is generally seen to provide the most representative measure of labour input because it takes into account the fact that workers are not homogenous, and as such, have different productivity levels (OECD, 2001). Because labour input, combined with capital, occupies the denominator of the productivity equation, information about its composition will improve productivity measures. In theory, the adjustment made to the LII can either raise or lower an estimate of multifactor productivity[3] (MFP), depending on how the composition of labour changes over time (Holmwood et al. 2005).

This paper provides a short summary of the findings of an investigation into the potential for the quality-adjustment of the New Zealand productivity series, serving two broad purposes. The first is to provide an outline of what quality-adjustment is in practice, drawing on economic theory and overseas case studies. The second purpose of the paper is to evaluate how conducive the labour data available in New Zealand is for quality-adjustment and to provide a proposal for the most suitable method for quality-adjusting the LII.

Why Adjust for Quality?

Reliable productivity measures are currently under high demand. Economic growth is accepted as being achieved through increases in the quantity of inputs or through productivity growth. Raising productivity is a vital move if New Zealand is to improve its economic performance; hence a representative measure of the variable is of high importance. Quality-adjustment is a significant move in improving our understanding of productivity in New Zealand.

To understand why we would quality-adjust the LII, we need to have a clear definition of what 'labour input' actually is. In practice, labour input should reflect the "Time, effort and skills of the work force" (OECD, 2001). Labour input data is generally based on hourly units capturing the time dimension, but not the skill dimension of labour. A quality-adjusted measure of labour input does not treat workers as a single input, but rather workers *of the same traits* are treated as the same category of input in the measurement of labour inputs. This cross-categorisation makes intuitive sense because in reality, workers of different skill-levels are not perfect substitutes, and firms treat them as distinct inputs in their production process (BLS, 1993).

To account for differences in quality, more weight should be given to units of relatively higher skill in aggregating units of labour input. In doing this, changes in labour composition that affect output will be reflected in changes in labour input, not as a change in productivity. As the OECD (2001) note, an increase in the average quality of labour implies that a quality-adjusted measure will rise faster than an unadjusted measure.

The fact that Multi Factor Productivity (MFP) is measured residually is widely considered as a concern, because there is uncertainty about the make-up of this term. Successful quality-adjustment effectively eliminates one of the possible components of the residual (OECD, 2001). As such, quality-adjustment should raise the confidence in any estimate of MFP, which will in turn raise the confidence of any results from research that is undertaken that uses the series as a data source.

The other major benefit of quality-adjustment is that the comparison of unadjusted and adjusted estimates of labour input can provide insight on how certain compositional changes affect productivity over time. In this sense, quality-adjustment can be seen as a step towards measuring the effects of "intangible investment" in the workforce (OECD, 2001).

Defining Quality

Put simply, a Quality-Adjusted Labour Index (QALI) uses proxies for skill to weight different categories of labour input. These different categories should include workers with similar skill sets. Higher weights are given to the growth rates of groups that earn higher wages, implying that higher wages represent higher marginal productivity and higher quality (Statistics Canada, 2003).

A person's quality is unobservable and is therefore very hard to measure. Although there has been substantial research on quality-adjustment, one of the main points of contention is whether we can accurately capture what quality really is.

Quality-adjustment requires a means to apportion workers into different categories, reflecting different skill sets. To do this, observable variables are used to attempt to cross-categorise workers into uniformly skilled groups.

The main variables that are accepted as representing the quality or skill set of a worker are their investments into education and training. As the ABS (2005) note, human capital theory stipulates that there is a positive relationship between wage levels and education attainment and work experience. Education and training that improve workers' skills and productivity, can be treated as investments in human capital.

The Use of Wages

In addition to portioning out workers according to skill sets, it is also necessary to value each of the groups to give an indicator of what their relative quality actually is. Internationally, wages play an integral part in quality-

adjustment, being the best available, albeit imperfect, variable for weighting groups in quality-adjustment. As Statistics Canada (2003) note, the theory of the firm specifies that, under certain conditions (the firm is a price-taker on labour markets and aims at maximising profits), labour of a particular type will be hired up to the point that the cost of an additional hour of labour is equal to the additional revenue that using the labour generates. Once each type is given a weighting, based on its share of the total labour costs (wages), a quality-adjusted LII can be compiled.

Of course, it would be naive to assume that wages are determined solely by a person's productivity. Other factors, such as discrimination, union bargaining, signalling and mismatch can compromise the validity of using wages of a measure of worker productivity (Schwerdt and Turunen, 2006)[4]. As overseas case studies have shown, there are measures that can be taken to control for some of these factors.

The Current New Zealand Series

Statistics New Zealand's current Labour Input Index (LII) is derived using a chained Törnqvist index in which weights are based on industry shares of the measured sector nominal labour income (including an estimate of the labour income of the self-employed). If we assume that inter-industry wage differentials reflect differences in skills, this goes some way towards quality-adjusting the series, because the weights will be comparatively large for industries that pay above-average wages and vice-versa for industries that pay below-average wages (Zheng, 2005). This adjustment, however, has some major limitations.

Firstly, it involves a reasonably crude portioning variable. Although some industries require more skill and are thus paid higher, the industry definitions are not particularly specific to the qualities of workers and hence include a variety of different 'skill levels'. This is because the correlation between industry and skill is far lower than for other variables, such as education and skill.

The other major disadvantage of this approach is that, from a policy perspective, it provides little insight into how labour composition affects aggregate productivity. A LII that is adjusted for education levels, for example, should provide some valuable information about issues such as how an increased number of people having tertiary qualifications affect productivity over time.

In addition to this point, the BLS (1993) note that wage differentials between industries can reflect allocative inefficiencies rather than differences in marginal products.

Methods of Cross-Classifying Workers

Before a quality-adjusted LII is created, workers first need be partitioned into distinct groups, based on the variables of interest. This section provides an outline of

the core characteristics of two general approaches to cross-classify labour: the 'Jorgenson' method (a.k.a. the 'Average Wage Model') that was developed by Jorgenson, Gollog and Fraumeni in 1987, and the 'BLS Method' (a.k.a. the 'Wage Model') which was developed by the US Bureau of Labour Statistics in 1993.

In theory, the two approaches should provide similar results, because the only real difference between them is how the total wages for each group is determined. Although both methods have their pros and cons, it is worthwhile attempting both approaches to ensure that that their results are not at odds.

The Jorgenson approach is used in the official measure of quality-adjustment in the UK, Canada and Denmark. The BLS Method is used as the preferred measure in the US and Australia, for the reasons explained below. Although the BLS method is conceptually preferable, the choice of method adopted appears to be largely driven by the data constraints faced by each country.

The Jorgenson Method

This first approach was adopted by Jorgenson et. al in 1987. The premise behind the Jorgenson technique is that individuals in the workforce can be cross-classified by a series of variables, with an average wage being taken for each category of worker. Once these wages are calculated, they are combined with information on the total number of hours worked for each category, to estimate each group's total wages. After this is summed across all categories, an estimate of total weighted hours is obtained, which can be indexed (Riley *et al.* 2005).

A primary advantage of this method over the BLS method is that it is comparatively simple, both conceptually and methodologically. The original approach used up to six characteristics (age, education, class of workers, gender, occupation, and industry) to cross-classify labour (Statistics Canada, 2003). The use of such a wide-ranging set of traits, however, can lead to difficulties in identifying sources of labour composition growth (BLS, 1993).

In its implementation, the main problem with using the Jorgenson method is that the partitioning of the population into subgroups requires a large sample size, otherwise it results in categories having not enough units for it to provide reliable weightings (BLS, 1993).

The BLS Method

This approach involves econometric regression analysis, in which a wage rate is used as the dependent variable, and the relevant indicators of skills are used as explanatory variables. The calculated earnings functions provide estimates of payments per hour associated with worker characteristics. The wages for each category are then used to provide weightings in the Törnqvist index.

As the BLS (1993) note, a benefit of this approach is that the nature of a regression allows for a more specific understanding of the causes of increases in productivity,

so long as there are not too many variables introduced in the regression. The analysis of the parameter values also allows for checking that the results are consistent with economic theory. As noted, in human capital theory, skills are the ultimate source of worker productivity, and education and training are the means for acquiring skills.

The other main benefit that the BLS (1993) suggest of using a wage equation is that it can increase the accuracy of the earnings estimate whenever earnings of different types of workers are structurally related. This makes this method superior in cases where we have categories of workers with few units in them, which would otherwise lead to high variance. This is one reason why the ABS prefer this approach, due to their data source for quality-adjustment being based on a reasonably small sample.

Additionally, if the data is available, this method allows for controlling of other factors, such as sex, region, and numbers of dependent children, that will affect the dependent variable (wages) but do not necessarily reflect a person's productivity. In this sense, the BLS provides a more representative measure.

Due to the benefits listed above, the BLS method widely seen as being the more useful of the two approaches. As mentioned, the main advantage of the Jorgenson method is that it is simpler, and of little disadvantage if a large dataset is available and if the additional information from regression analysis is not of major interest. Statistics Canada has applied both methods, but is moving towards just using the Jorgenson method, largely because have provided very similar results and it is easier to construct.

Constructing Weights

Once workers have been appropriately cross-classified, weights are calculated by putting each category's shares of wages into a Tornqvist formula to create an index number for each period.

The ABS use the following equation, based on the difference of labour inputs of successive periods. These are weighted by each category's average shares of labour compensation over two periods.

$$\Delta \ln L = \sum_j \frac{1}{2}\left(s_{lj}(t) + s_{lj}(t-1)\right)\Delta \ln h_j$$

Where $\Delta \ln L$ represents the first order difference of logarithms of aggregate labour input, $\Delta \ln h_j$ represents the first order difference of logarithms of the ith type of labour input, $s_{lj}(t)$ and $s_{lj}(t-1)$ are the value shares[5] of the jth labour inputs in the total labour costs in periods t and t-1.

Data Requirements

Because official productivity statistics are a reasonably new phenomenon in many countries, the series are invariably at the mercy of what appropriate datasets have

previously been compiled, often leading to ad-hoc methodologies, and less-than-ideal proxies being adopted.

The OECD describes quality-adjustment as "Desirable but difficult", stressing its potential payoffs, but noting that explicit differentiation is both data and research intensive.

In their 2001 Productivity Manual, the OECD note

> "As a minimum, a time series of hours worked, broken down by one differentiating characteristic, has to be available, alongside corresponding statistics for average compensation, broken down by the same characteristic."

The OECD further add

> "Even when only a simple trait such as occupation is chosen to differentiate labour input, information requirements are severe: data are needed that distribute the number of total hours worked across different occupations, by individual industry and by individual year. In addition, quantity measures of labour input (hours) have to be accompanied by price measures relative average compensation to construct weights for aggregation."

As the New Zealand application that follows shows, although quality-adjustment is definitely feasible, New Zealand is no exception to the points mentioned above as there are a number of challenges that have to be overcome with the data.

Using New Zealand Data

There are a number of sources of labour data in New Zealand. The most suitable data sources for quality-adjustment in New Zealand are Linked Employee-Employer Data (LEED), the Household Labour Force Survey (HLFS) and its annual supplement, the New Zealand Income Survey (NZIS), as well as the Census of Population and Dwellings. In the context of quality-adjustment, all of the data sources have various pros and cons, so there are inevitable tradeoffs for any choice of dataset.

Linked Employee-Employer Data (LEED)

The current LII uses data sourced from the QES and the Business Demography database, supplemented by data from the Population Census and the HLFS. LEED data is planned to be introduced as the main input in the near future. There is obvious merit in using the same data sources for the adjusted and unadjusted indices, however other factors need to be considered when choosing a data source.

In many ways, LEED is a very suitable data source for quality-adjustment. It includes longitudinal information on income, age, sex, industry, experience, tenure for virtually the entire working population, dating back to 1999. In the future, LEED may also be linked to information on educational attainment from Ministry of Education (MOE) data, which would be vital if LEED is to be applied in this context.

Despite its strengths, LEED's limited back series is a concern for a number of reasons. Firstly, for a productivity series to be of any great interpretative value, it should span a reasonable time length, so that business cycles can be isolated.

A more specific issue with the short back series is that it limits the use of LEED's experience and tenure variables, as at present we will not be able to cross-classify according to many levels of either variable. This, of course will be less of an issue as time progresses. In comparison to the other data sources, this characteristic of the series is not such of an issue. The longitudinal nature of LEED allows for information on the actual experience and tenure (subject to the back series constraints), which is something that no other New Zealand series can give, and is rare internationally. Additionally, the need to have a large back series for these variables (e.g. up to 30 years) should not be stressed too far, due to the diminishing returns that are inherent to both experience and tenure.

Experience is used for in quality-adjustment as a proxy for training, but even experience is often estimated. The ABS calculate 'Potential Experience' which is given as a person's age, less the number of years for preschool, the estimated number of years in education, with one year per child subtracted for women with up to four children. The UK and Canada simply use age for experience, which is provided in LEED.

Although education data may be linked to LEED in the future, its application to quality-adjustment may add biases to the series. The MOE data would provide information on educational attainment from 1997 onwards. This will cause some major issues, as any education achieved before this period will not be recognised. In terms of cross classifying the population, we cannot distinguish between people who have obtained qualifications prior to this period, and those that have not attained any at any point. As for experience, this limitation will diminish as time goes on, but it would be some time before we could consider the inaccuracies to be negligible. Additionally, education achieved overseas would be omitted.

The other major limitation of LEED data is that it does not provide information on hours worked. Labour input is measured in terms of hours, so again, this will be a limitation in this context. A possible way to partially offset this problem would be to use the average hours worked per industry from the Quarterly Employment Survey, and apply this to each unit in the industries.

Despite its current limitations, LEED includes a lot of information that could be used for improving our current 'partial' quality-adjustment by industry. Experience (or age), tenure, industry, region, and sex could all be used to cross-classify the series. Although it is far from a perfect proxy, there will be some degree of correlation between industry and education because some industries demand higher qualifications than others. Taking this approach could improve our LII estimates, however it could not be considered as providing a quality-adjusted series in the internationally comparable sense.

Household Labour Force Survey (HLFS)

The HLFS provides potentially useable information on workers' age and number of dependent children (which could be used to estimate 'Potential Experience'), education, gender, occupation, industry, hours worked, and income (which comes from the annual NZIS). The combined data runs back to 1997, providing a slightly longer time series than LEED. Although the back series is limited, the surveys have changed little over time, so they lend well to time series applications. In addition, the limited back series does not lead to the same complications as it does for LEED.

Compared to other possible data sources, its main disadvantage is that it is sample-based. The HLFS surveys 30,000 individuals in the civilian non-institutionalised usually resident population aged fifteen years and over. While this is a large sample, when compared to LEED or Census datasets, it is very small. The response rate for the HLFS is generally around 90 percent, while the response rate for the income supplement is around 85 percent of those that respond to the HLFS. Both of these factors lead to sample error.

An additional issue with the small sample size is that it limits the cross-classifying options. If the Jorgenson approach is adopted, using the HLFS would result in some groups having very few units, leading to large variability and potentially unreliable estimates. To counter this, less distinguishing variables could be used, or less specific divisors, but both of these are undesirable as there will be more heterogeneity within groups.

Another disadvantage of the HLFS is that it is not a major source for the current LII. Obviously, it would be ideal to use the same data source for both the original and quality-adjusted productivity series. A simple quality-adjusted series could be compiled by using HLFS and cross-classifying it accordingly. The issues with this are firstly, the HLFS is not seen as the best source for productivity estimates, and secondly, comparisons between this and the original series casts questions over whether differences are due to the adjustment, or rather to the different data sources.

Another issue with the HLFS is that, unlike LEED, it is not based on administrative data, which makes it subject to more non-sampling error in some senses. The variables chosen are subject to respondent error, making the

education and industry categories not as reliable as they could be.

Census of Population and Dwellings

Like the HLFS, the Census includes reasonable proxies for all of the main variables that would be required for quality-adjustment. A benefit of the Census is that it has by far the longest back series, which is very important for analysis of productivity over business cycles. It would be ideal to link Census data to LEED data at the unit-record level, however this would be a very large undertaking and may not be feasible due to privacy reasons.

The main disadvantage of the Census in the context of quality-adjustment is its frequency. The five year gap between Censuses causes two issues. Firstly, the information will not be very timely if we require it for recent productivity estimates. For example, if we were to use the Census to quality-adjust our 2006 estimates we would have to use data from 2001 to estimate of our variables, because the information from most recent the Census would not be available. The other issue is that the five year gap between data points casts question over what happened to the variables of interest in these periods.

Another problem with the application of the Census for a time series is that its questions and data processing levels have changed considerably between Censuses, leading to problems with comparisons of data points. Although this doesn't affect all the variables substantially, the education question has changed significantly between 1996 and 2001, which appears to have affected responses, making the results reasonably inconsistent. This is one issue that has to be monitored if the Census is used.

The other issue with using the Census is that its income information poses two problems for quality-adjustment. Firstly, the questionnaire asks for income in bands, which means that exact estimates of people's incomes are not available. Secondly, the Census asks for a person's total income, which could include income from any number of sources –not conceptually ideal for quality-adjustment.

Interpolating the Census and HLFS

Despite its undesirable frequency, the Census could potentially be used for quality-adjustment if the data was smoothed. This could be done by simple methods, such as having linear trends between data points for the series, however a better method would be to interpolate it with HLFS data. This could be done by using the Census as a benchmark, and obtaining information on groups from the HLFS (e.g. industry), applying it to all the (more specific) industry groups in the Census. Because turning points are very important for quality adjustment, the interpolation will be necessary if we want an accurate series.

The combination of Census and HLFS would provide information on potential experience (for which we can use the ABS formula) and education (although there are

issues with comparability, but we could use just a few broad categories as the ABS do).

Additionally, sex, industry and region could also be incorporated. The inclusion of hours in the series allows for to inference into whether a person works part-time or full-time.

Interpolated series have been used for quality-adjustment in both Australia and Canada, suggesting that this move would not make this series inferior by international standards. As the ABS stress, there are inevitable complications with interpolating data to this degree of detail.

Conclusions and Recommendations

This paper has aimed to provide background into quality adjustment, and to summarise the feasibility analysis that has been undertaken. As the paper has argued, quality-adjusting the LII would provide a very desirable enhancement to Statistics New Zealand's current productivity series. The main benefit of the enhancement is that it would provide a more representative measure of labour input; however it can also provide additional information about how the compositional change in the labour force affects productivity.

When conducting quality-adjustment, there are a number of key decisions that have to be made. Although there is little debate about how the Tornqvist index is compiled internationally, the cross-categorisation procedures that precede this step differ markedly between countries. Although the BLS method has a number of benefits for interpretative puposes, it is not used as the prevailing approach in every country. The decision of which method to use appears to be partly dependent on whether the additional information that the BLS approach offers is of particular interest to stakeholders, but it is largely contingent on what data sources are available to each study.

As is often the case with productivity research, finding suitable data is the greatest challenge. Analysts need to find an approach that maximises the worth of the information at hand, without stretching the application of the data too far. Common data issues that this paper has highlighted are that adequate proxies for training are hard to obtain, wages are seen as an imperfect proxy for productivity, and sample sizes are often too small to adequately apply the Jorgenson method.

To apply an ideal quality-adjustment, a large sample that incorporates all of the variables desired for cross classification, as well as income and hours worked, would be needed. In addition to this, if the series for creating weights is not the same as that used in the original LII, there needs to be some means to link the two series.

In terms of an application to New Zealand, this is where one of the main challenge lies. The LII will soon be based on LEED, which will be an excellent source of future data for quality adjustment in most respects, however its lack of information about educational attainment is a key weakness in this context. Because education is generally seen as the most important variable to include for cross-classification, this makes LEED unusable for an internationally comparable quality-adjustment in the near future.

Assuming that education data is added to LEED, it will unambiguously be the best data source for quality-adjustment in the future. This is largely due to the longitudinal information that it can provide, as well as additional information that comes from it having data from both the supply and demand sides of the labour market.

In terms of the potential for education data to be added to LEED, the need to have a long back series is far more important. As time goes on, the percentage of the workforce that obtained qualifications before 1997 will decrease, however this group will be very substantial at present, and it will be a long time before the misclassifications that can result from this group will be negligible. Modelling education using one of the other data sources could be attempted, but this would be complicated, and it is not a desirable step to have to take for such a key variable.

The HLFS/ NZIS and Census both provide information that would be adequate for quality-adjustment, however combining them through benchmarking the HLFS movements to the five-yearly Census totals would minimise some of the weaknesses of the two sources. To quality-adjust using these sources, we could either use the interpolated series as the series in which the weights can be applied to, or take the alternative approach of applying the weights to the original LEED LII series.

At present, the interpolation between the HLFS/ NZIS with Census would be the most viable option for quality-adjustment. In the future, LEED will supersede these series as the best data source for quality-adjustment, however this will not happen in the next few years, particularly if education data is not included.

Further Research

The development of a quality-adjusted measure of productivity is planned to begin in 2007. An experimental quality-adjusted series should follow.

Notes

1   Value-added is defined as gross output less the value of intermediate inputs.

2   Because this paper is intended to be a stand-alone piece, the background to the current labour series that is provided in this paper is limited to what is relevant to this project. For more detailed information about Statistics New Zealand's current productivity series, consult, www.stats.govt.nz/developments/productivity.htm

3   Statistics New Zealand currently measure multifactor productivity, as opposed to single factor or total factor productivity. Strictly speaking, multifactor productivity is a more narrow concept than total factor productivity as it only includes capital and labour as inputs.

4   Additionally, wage rates are often seen as being partly determined by differences in firm behaviour. This factor is not considered in this paper, however there is currently research being undertaken on LEED data to investigate the relevance of this factor.

5   Obviously, the shares for all groups of workers should sum to one each year.

References

Bureau of Labor Statistics (1993). *Labor Composition and U.S. Productivity Growth, 1948-90*, U.S. Department of Labor, Bureau of Labour Statistics, Bulletin 2426, Washington DC: U.S. Government Printing Office.

Organisation for Economic Co-operation and Development (2001). *OECD Manual: Measuring Productivity: Measurement of Aggregate and Industry-Level Productivity Growth*, France: OECD.

Reilly, R., Milne, W. and Zhao, S. (2005). *Quality-Adjusted Labour Inputs*. Australian Bureau of Statistics Research Paper No 1351.0.55.010.

Holmwood, R., Lau, E., Richardson, C. and Wallis, G. (2006). Experimental quality-adjusted labour input measure – An update. *Economic Trends 631 June 2006*.

Statistics Canada (2003). *A Revision of Statistics Canada's Estimates of Labour Input for the Productivity Accounts*. Statistics Canada.

Statistics New Zealand (2006). *Productivity Statistics: Sources and Methods*. Wellington: Statistics New Zealand.

Zheng, S. (2005). *Estimating Industry-Level Multifactor Productivity for the Market-Sector Industries in Australia: Methods and Experimental Results*. Australian Bureau of Statistics Research Paper No 1351.0.55.004.

Author

Thomas McNaughton
Statistical Analyst
National Accounts
Statistics New Zealand
P.O. Box 2922
Wellington
Thomas.Mcnaughton@stats.govt.nz