Credit Where Credit Is Due: An Approach to Education Returns Based on Shapley Values

Barakat, Bilal; Crespo Cuaresma, Jesus

DOI:
10.57938/611b510d-bae8-4c76-9a07-0ab284411596

Published: 01/06/2016

Document Version
Publisher's PDF, also known as Version of record

Department of Economics
Working Paper No. 228

Credit Where Credit Is Due: An Approach to Education Returns Based on Shapley Values

Bilal Barakat
Jesús Crespo Cuaresma

June 2016

WIRTSCHAFTS
UNIVERSITÄT
WIEN VIENNA
UNIVERSITY OF
ECONOMICS
AND BUSINESS

Credit Where Credit Is Due: An Approach to Education Returns Based on Shapley Values

Bilal Barakat[a,b] *(corresponding author), bilal.barakat@oeaw.ac.at*

Jesus Crespo Cuaresma[a,c,d]*, jcrespo@wu.ac.at*

[a] *Wittgenstein Centre for Demography and Global Human Capital*

[b] *Vienna Institute of Demography (VID), Austrian Academy of Sciences, Welthandelsplatz 2, 1020 Vienna, Austria*

[c] *Vienna University of Economics and Business (WU), Welthandelsplatz 1, 1020 Vienna, Austria*

[d] *Austrian Institute of Economic Research (WIFO), Arsenal Objekt 20, 1030 Vienna, Austria*

Abstract

We propose the use of methods based on the Shapley value to assess the fact that private returns to lower levels of educational attainment should incorporate prospective returns from higher attainment levels, since achieving primary education is a necessary condition to enter secondary and tertiary educational levels. We apply the proposed adjustment to a global dataset of private returns to different educational attainment levels and find that the corrected returns to education imply a large shift of returns from tertiary to primary schooling in countries at all income levels.

Keywords: Returns to education, Shapley value.

JEL codes: I26, I25, C71.

1 Introduction

Numerous studies have aimed at assessing the relationship between education and earnings and obtaining estimates of the private returns to different attainment levels. Typically, such estimates derive from earnings regressions and are interpreted as the expected premium for an extra year of schooling in a particular attainment level, keeping other personal characteristics constant. This equates the return to primary, say, with the earnings premium of primary graduates only, thereby ignoring that '*primary education is a necessary input into further levels of education which may have higher economic returns. If the benefit that primary education confers by permitting access to more lucrative levels of education is taken into account, its true return will increase.*' (Colclough, Kingdon, and Patrinos 2009, 3)

To our knowledge, only Appleton, Hoddinott, and Knight (1996) have investigated this issue quantitatively. They include in the return to primary the *expected* prospective returns of secondary and tertiary education (that is, weighted by the probabilities of subsequently completing those cycles), and demonstrate that taking the sequential nature of attainment progression seriously can greatly increase the estimated primary return. However, this approach cannot be applied simultaneously to returns at all attainment levels without double-counting.

Here, we use an analogy between the problem of (re-)attributing returns between primary, secondary and tertiary education and cost allocation based on the Shapley value. Such a reinterpretation of the problem yields returns at *all* levels of schooling that account for their prerequisite relationships. Using the global dataset on education returns recently developed by Montenegro and Patrinos (2014), we calculate adjusted returns for all countries with available data. Our results indicate that adjustments based on the Shapley value generate a large shift of returns towards primary in countries at all income levels, for both males and females.

2 The Shapley value and returns to education

The Shapley value provides a conceptual basis to the distribution of coalition pay-offs with desirable properties. Specifically, only this division rule satisfies additivity, efficiency (the full coalition pay-off is distributed), symmetry (equal pay-off for substitutable players) and the so-called null-player axiom (players not contributing to the pay-off of any coalition receive zero pay-off). While attainment levels are not perfectly comparable to players in a coalition game, the logic of Shapley values of 'fair shares' derived from the expected marginal contribution of players entering a coalition makes it suitable for the computation of prospective returns to education.

Assuming a strictly linear dependent ordering of successive schooling levels $i \in \{0, 1, \ldots, N\}$, the Shapley approach yields particularly simple results. Denote by r_i the marginal return to attainment level i over level $i-1$. When derived from coefficients in a log-earnings regressions, these returns are in units of the baseline wage for an individual with no education.

The Shapley value corresponds to the sum

$$\frac{1}{|N|!} \sum_{R} \left[v(P_i^R \cup i) - v(P_i^R) \right],$$

over all possible orderings R of the N players, where P_i^R is the set of players preceding player i in R, and $v(S)$ is the worth of coalition S. From here on, we assume $N = 3$.

In the context of returns to different edcational attainment levels, this leads to payoffs shown in Table 1. For instance, in the second row, when the primary-education 'player' joins the empty coalition, its worth increases by the primary wage premium r_P. When the tertiary-education 'player' joins next, this does not increase the coalition's worth, since exploiting the returns of tertiary education depends on (yet absent) secondary. The wage only jumps from a primary-graduate wage to a tertiary-graduate wage when the secondary-education 'player' joins. Summed across orderings, the tertiary marginal wage premium r_T is shared equally between all 'players', r_S is shared equally between the primary and secondary 'players', and the primary 'player' retains all of r_P.

order	primary	secondary	tertiary
PST	r_P	r_S	r_T
PTS	r_P	$r_S + r_T$	-
SPT	$r_P + r_S$	-	r_T
STP	$r_P + r_S + r_T$	-	-
TPS	r_P	$r_S + r_T$	-
TSP	$r_P + r_S + r_T$	-	-

Note: $r_P, r_S,$ and r_T represent the marginal wage premiums.

Table 1: Marginal payoffs to primary/secondary/tertiary education 'players' as they join the coalition in every possible sequence.

Labeling the fair share of the marginal return to attainment level j that is attributable to level i as r_{ij}, the

above derivation generalises to the straightforward expression

$$r_{ij} = \begin{cases} r_j/j & \text{, if} \quad i \leq j, \\ 0 & \text{otherwise.} \end{cases}$$

Accounting for all returns requires *combining* the return to primary education for those with primary only with the returns attributed to primary for those with higher attainment. Let p_i be the share of the population that completed level i. The conditional probabilities q_{ji} of having completed at least level j among those having completed at least level i are: $q_{ji} = \sum_{n=j}^{N} p_n / \sum_{n=i}^{N} p_n$ for $j \geq i$. The adjusted return is simply the weighted sum $\tilde{r}_i = \sum_j r_{ij} q_{ji}$.

Alternatively, consider $\tilde{r}_i^- = r_{ii}$, the return attributed to level i for an individual with attainment i. This has the advantage of not requiring knowledge of the participation rates. However, by reducing the returns to higher levels without increasing the returns at lower levels, some returns remain un-attributed. We therefore refer to these as 'decremental' adjusted returns, indicated by the minus-sign superscript.

Both adjusted returns are comparable to conventional marginal returns in that they capture the benefit attributed exclusively to a given attainment level, net of the contribution of lower levels. They are, however, *pseudo*-marginal in that they do not relate directly to an individual's marginal choice of continuing or not continuing education to the next level. For example, a primary graduate's choice to continue with secondary education will 'retrospectively' change their return to primary.

Note that the adjusted returns do not decompose with respect to nested education levels. This reflects a general property of the underlying Shapley values. In particular, adjusted returns for individual years of schooling do not aggregate up to the adjusted returns for whole attainment levels shown here. Arguably, this is less of a limitation than it might seem. While conventional estimates are commonly expressed as returns to a single year of schooling, this is most often strictly a matter of scaling, and does not reflect the actual estimation. Regardless of whether results are shown at the per-year scale (as below for consistency with the published conventional estimates), the adjustment itself is performed on whole-level returns.

3 Empirical application: Revisiting returns to education

3.1 Data

Montenegro and Patrinos (2014) present homogenized estimates of returns to schooling, by sex, for 139 countries. The adjusted returns additionally require the proportion completing each attainment level. These are extracted from the World Bank's World Development Indicators (World Bank 2013), specifically the 2010 values from the Lutz, Butz, and KC (2014) time series on attainment of the population aged 25 and above, which covers more countries than the Barro-Lee data (Barro and Lee 2013).[1]

3.2 Adjusted returns to education: A comparison

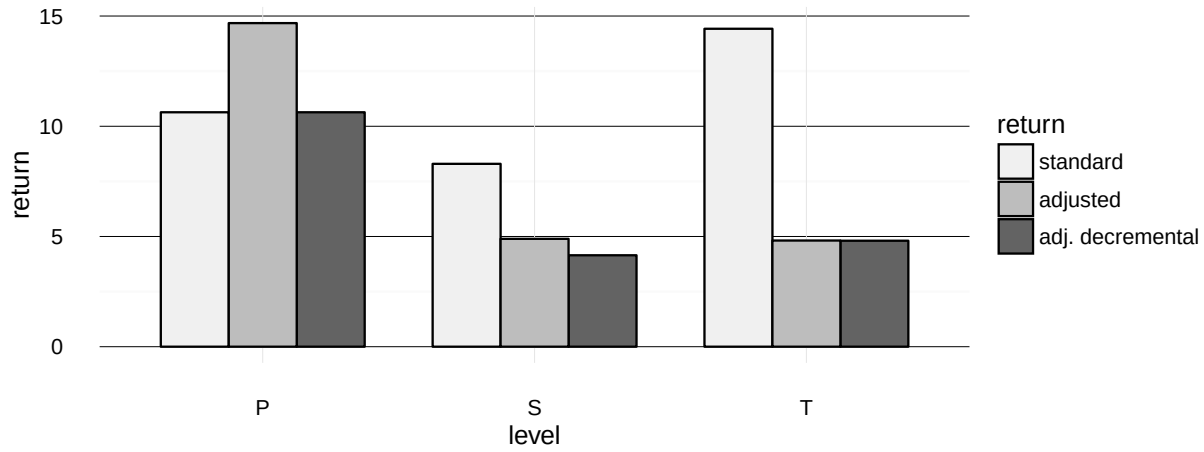

Figure 1: Average returns by adjustment type: unweighted world sample

Using the above data, we compute adjusted returns to education for three attainment levels and both sexes. Unweighted averages aggregated over sex and all countries are depicted in Figure 1, annualised to match the scale of the standard figures.

Figure 1 shows that the position of primary and tertiary clearly flip after the adjustment, with adjusted primary returns far exceeding adjusted secondary and tertiary returns, which are approximately equal. Returns to secondary education are affected both positively and negatively by the re-attribution implied by the adjustment proposed, with a negative net effect reflecting the relatively low participation in, and therefore gains flowing down from, tertiary education.

[1] Our conclusions remain entirely unchanged if the Barro-Lee data are used instead. Those results are available from the authors on request.

The large reduction in the return attributed to tertiary education implied by the proposed adjustments is not balanced by the increasing returns to primary and secondary. For the decremental adjustment, this is unsurprising, given its 'leaky' definition. For the preferred balanced adjustment, which does re-attribute all returns, this is explained by the fact that the *adjusted* returns to primary, for instance, accrue to *all* individuals with at least, not just the ones with exactly, primary education attainment.

Disaggregation by gender and country income level (Figures 2) confirms that the effect of the adjustment is most accurately described as a shift *towards primary* (as opposed to a shift *away from tertiary*). There are no noticeable gender differences with respect to the effect of the adjustment. It is also evident that the bulk of the shift results from the re-attribution away from the higher levels, rather than the re-attribution towards primary. As a result, the purely decremental adjustment, which can be calculated even without knowledge of participation rates, represents a surprisingly good approximation to the full adjustment in terms of indicating the overall shift in balance.

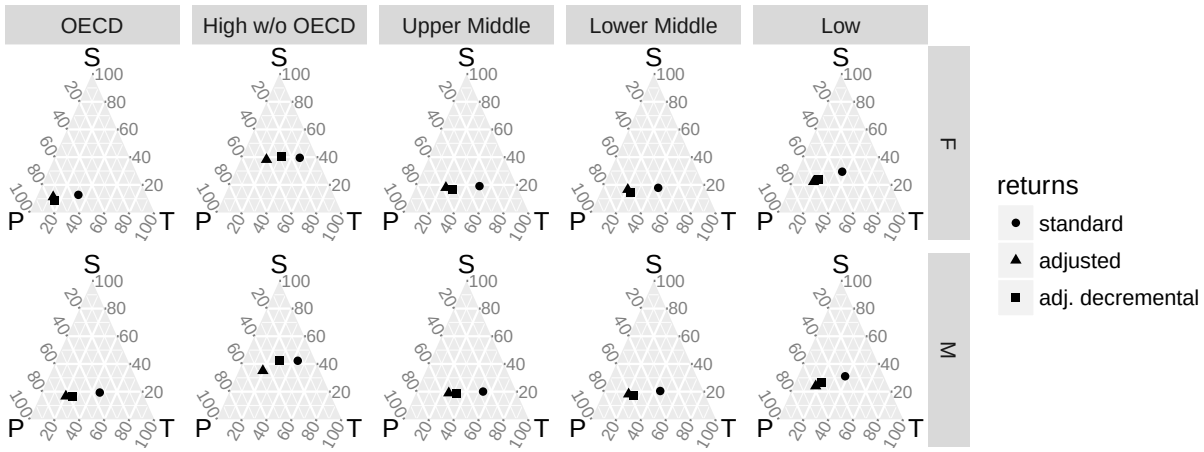

Figure 2: Relative magnitude of returns relative to baseline by gender and income group

4 Discussion and policy implications

We propose a method based on Shapley values to adjust returns to education across attainment levels, thus accounting for the prospective nature of lower educational attainment levels as necessary conditions to achieve higher ones. While theoretically justifiable, the magnitude of the adjustment implied by Shapley values may seem excessive. Letting the tertiary level only 'keep' one third of its conventional returns is certainly a large adjustment. The result that this adjustment is sufficient to more than *reverse* the positions of primary and tertiary education with respect to their marginal returns suggests that the 'true' marginal returns to lower

and higher levels of schooling may be aprroximately equal if we go some way, but not quite as far, in the direction of recognising the dependence of the latter on the former.

As noted, the Shapley-adjusted returns are pseudo-marginal in that they do not directly relate to the individual decision problem of educational progression. While individuals need not care about which education level their earnings are 'attributed' to, this question is, however, of great interest for setting public policy. From this perspective, the difference in interpretation between 'proper' marginal and pseudo-marginal returns is not clear-cut, however. Neither refer to the use of 'marginal dollars of public funds', after all. Even conventional marginal returns are therefore correctly interpreted as merely motivating general development priorities. Our conclusion, therefore, is that if dependencies between school levels are taken into account, prevailing conventional 'face value' returns to different levels of education do not differ sufficiently to serve as a sound justification for strongly privileging one level over another in prioritising educational development.

Acknowledgements

The authors would like to thank Felix Mendoza for research assistance.

References

Appleton, Simon Mark, John Hoddinott, and John B Knight. 1996. "Primary Education as an Input into Post-Primary Education: A Neglected Benefit." *Oxford Bulletin of Economics and Statistics* 58 (1). Department of Economics, University of Oxford: 211–19.

Barro, Robert J, and Jong Wha Lee. 2013. "A New Data Set of Educational Attainment in the World, 1950–2010." *Journal of Development Economics* 104: 184–98.

Colclough, Christopher, Geeta Kingdon, and Harry Anthony Patrinos. 2009. "The Pattern of Returns to Education and Its Implications." *Policy Brief 4, Research Consortium on Educational Outcomes and Poverty*.

Lutz, Wolfgang, William P Butz, and Samir KC. 2014. *World Population and Human Capital in the Twenty-First Century*. Oxford University Press, Oxford (UK).

Montenegro, Claudio E, and Harry A Patrinos. 2014. "Comparable Estimates of Returns to Schooling Around the World." *World Bank Policy Research Working Paper*, 7020.

World Bank. 2013. *World Development Indicators 2013*. The World Bank.