EJW
ECON JOURNAL WATCH
Scholarly Comments on Academic Economics

ECON JOURNAL WATCH 12(3)
September 2015: 339–345

Education Premiums in Cambodia: Dummy Variables Revisited and Recent Data

John Humphreys[1]

LINK TO ABSTRACT

In their 2010 *Asian Economic Journal* paper, Ashish Lall and Chris Sakellariou made a valuable contribution to the understanding of education in Cambodia. Their paper represents the most robust analysis of the Cambodian education premium yet published, reporting premiums for men and women from three different time periods (1997, 2004, 2007), including a series of control variables in their regressions, and using both OLS and IV methodology.[2]

Following a convention of education economics, Lall and Sakellariou (2010) use a variation of the standard Mincer model (see Heckman et al. 2005), where the logarithm of wages is determined by education, experience, and a series of control variables, including locality (urban/rural), employment sector (public/private) and marriage status. Lall and Sakellariou report results separately for males and females.

Since the Mincer model is log-linear, the coefficient for total years of education—a continuous variable—can be interpreted, without any need for adjustment, as the education premium. For the 2007 data, for example, Lall and Sakellariou find an education coefficient of 0.066 for men and 0.068 for women (2010, 342), which means that every additional year of schooling is expected to

1. Professional Research Institute for Management and Economics, Phnom Penh 12304, Cambodia.

2. The reader should note that Lall and Sakellariou (2010) calculated the 'education premium' and not the 'return on education.' While related, these two concepts are not the same: the education premium is a measure of the change in income associated with a change in education, whereas the return on education measures the relationship between the education premium and the cost of attaining that education. Lall and Sakellariou are clear and consistent throughout their paper in focusing on the education premium, and consequently in this paper I likewise concentrate solely on the education premium.

increase incomes by 6.6 and 6.8 percent respectively. I do not challenge the above estimates, which remain the best published estimates for the Cambodian education premium.

Continuing with convention, Lall and Sakellariou also provide a model that allows for non-linearity in the education premium. They do this by replacing the continuous education regressor with a series of education dummy variables, one for each of four levels of education (primary, junior high, senior high, university). Based on the coefficients for these dummy variables, Lall and Sakellariou report an annualized education premium for each different level of education. Where the first model estimated, e.g., an education premium of 6.8 percent for women in 2007, Lall and Sakellariou reported that the non-linear model estimated an annualized education premium of 7.9 percent for primary school, 5.5 percent for junior high, 7.8 percent for senior high, and 12 percent for university (2010, 343).

Those estimates, however, are based on an error in how Lall and Sakellariou interpreted the dummy variable coefficients, which results in a downward bias. Robert Halvorsen and Raymond Palmquist (1980) noted that dummy variable coefficients in log-linear models are commonly misinterpreted, pointing to several examples including papers by luminaries such as Robert Lucas (1977) and Zvi Griliches (1971). At least Lall and Sakellariou are in good company.

Again, the coefficient for a continuous regressor can be interpreted directly to be the premium. But Halvorsen and Palmquist (1980) pointed out that, when the regressor is a dummy variable, the coefficient c needs to be transformed by $\exp(c)-1$ before it can be interpreted as a premium. Further, Peter Kennedy (1981) observed that the Halvorsen and Palmquist transformation requires that we know the value of c with certainty, and Kennedy provided a more robust solution where the estimated coefficient c' needs to be transformed by $\exp[c'-\frac{1}{2}\mathrm{Var}(c')]-1$ before it can be interpreted as a premium. Lall and Sakellariou did not use either of these transformations, and the omission substantially changes the results.

In correspondence, Chris Sakellariou kindly provided the coefficients and t-values for each of the dummy variables; I reproduce these as column (1) in Tables 1 and 2 below. Then, I report three different results for the premium: annualized premiums as calculated by Lall and Sakellariou (column 2); annualized premiums calculated by me using the Halvorsen and Palmquist transformation (column 3); and annualized premiums calculated by me using the Kennedy transformation (column 4).

TABLE 1. Annualized education qualifications premiums, among men ages 22–65 employed for wages

	(1) Coefficient (and t-value)	(2) Lall and Sakellariou (2010) result	(3) Halvorsen-Palmquist transformation	(4) Kennedy transformation
1997				
Primary	0.042 (0.6)	1.4%	1.4%	1.3%
Junior high	0.168 (2.2)	4.2%	4.7%	4.6%
Senior high	0.273 (3.3)	3.5%	4.4%	4.3%
University	0.573 (3.8)	7.5%	11.5%	11.1%
2003/04				
Primary	0.273 (4.7)	9.1%	10.5%	10.4%
Junior high	0.513 (8.8)	8.0%	11.9%	11.9%
Senior high	0.702 (10.9)	6.3%	11.6%	11.5%
University	1.186 (13.1)	12.1%	31.4%	31.2%
2007				
Primary	0.204 (1.8)	6.8%	7.5%	7.3%
Junior high	0.438 (4.5)	7.8%	10.8%	10.8%
Senior high	0.639 (6.3)	6.7%	11.5%	11.4%
University	1.091 (7.9)	11.3%	27.1%	26.6%

TABLE 2. Annualized education qualifications premiums, among women ages 22–65 employed for wages

	(1) Coefficient (and t-values)	(2) Lall and Sakellariou (2010) result	(3) Halvorsen-Palmquist transformation	(4) Kennedy transformation
1997				
Primary	0.009 (0.1)	0.3%	0.3%	0.2%
Junior high	0.066 (0.5)	1.9%	2.0%	1.8%
Senior high	−0.033 (0.2)	−3.3%*	−3.4%	−3.5%
University	−0.023 (0.2)	0.3%	0.2%	0.4%
2003/04				
Primary	0.219 (3.6)	7.3%	8.2%	8.1%
Junior high	0.438 (5.6)	7.3%	10.2%	10.1%
Senior high	0.669 (7.0)	7.7%	13.4%	13.3%
University	1.265 (7.1)	14.9%	39.8%	38.6%
2007				
Primary	0.237 (3.6)	7.9%	8.9%	8.8%
Junior high	0.402 (4.5)	5.5%	7.6%	7.5%
Senior high	0.636 (3.9)	7.8%	13.1%	12.5%
University	1.116 (6.8)	12.0%	29.1%	28.7%

*In their paper, Lall and Sakellariou reported this value as −1.1 percent (2010, 343, Table 8), based on the total senior high coefficient instead of the relevant difference.

To calculate the annualized premium for a given level (e.g., university), the process is to convert regression coefficients into premiums, then take the difference between the premium for the level of interest and the premium for the preceding level (e.g., the university premium *minus* the senior high premium), and then divide that difference by the number of years of study for the given level (e.g., four years for university).[3]

To provide a numeric example, the regression results for men in 2007 show a university coefficient of 1.091 and a high school coefficient of 0.639 (see bottom two rows of Table 1). But at this point Lall and Sakellariou did *not* transform the coefficients into premiums. They took the difference between the two untransformed coefficients (1.091−0.639 = 0.452) and then divided that by the average number of years at university (0.452/4 = 0.113), and so they reported an annualized premium for the university level of 11.3 percent. In contrast, if they had transformed the coefficients prior to subtracting and dividing, they would have found that annualized premium to be 27.1 percent (Halvorsen-Palmquist transformation) or 26.6 percent (Kennedy transformation).

As shown in the above tables, the adjusted results are significantly different from those originally published in Lall and Sakellariou, especially in the case of university education. The new results suggest that education premiums in Cambodia are much higher than previously thought.

The updated premiums are high by international standards, but the premiums build off a very low base, so the dollar increase is not large. Once the cost of university is factored in, the rate of return on education is less impressive, though still above average. Exploring the reasons for the relatively high university premium goes beyond the scope of this article, but one possible reason is that a university degree can help graduates from regional towns to get a job in Phnom Penh, where wages are higher.

A minor addition: More recent data

It is now possible to extend the work of Lall and Sakellariou (2010) by applying their model to more recent versions of their source data. In their original paper, Lall and Sakellariou applied a consistent model to data taken from the Cambodian Socio-Economic Survey (CSES) for 1997, 2003/04, and 2007. It may be of interest to provide the equivalent results based on the 2010 CSES results.

The 2010 CSES results provided by Cambodia's National Institute of Statistics did not include one of the variables used by Lall and Sakellariou (location

3. University takes four years; each other level takes three years.

= rural/urban) and so it was necessary to create a proxy variable (work type = agriculture/other) as a replacement. In another minor difference, the original Lall and Sakellariou model used "rural" as the base case and provided a coefficient for the dummy variable "urban,", while the model applied to the 2010 data uses "non-agriculture" as the base case and provides a coefficient for the dummy variable "agriculture." A third difference between the Lall and Sakellariou models and this imitation is that the former use only data for employed persons of ages 22 to 65, while I use data for employed persons of any age.[4] Unfortunately, these differences may make a direct comparison unreliable.

TABLE 3. Regression coefficients by gender, among persons of any age who are employed for wages, 2010 data

Regression:	(1)	(2)	(3)	(4)
Variable	Men	Women	Men	Women
Years of schooling	0.083 (18.1)	0.069 (12.1)		
Primary			0.161 (3.6)	0.102 (2.1)
Junior high			0.298 (5.7)	0.257 (4.2)
Senior high			0.709 (11.7)	0.685 (8.9)
University			1.301 (18.1)	1.220 (12.5)
Experience	0.024 (4.3)	0.007 (1.3)	0.021 (3.9)	0.009 (1.8)
Experience^2	−0.0003 (3.3)	−0.0001 (0.7)	−0.0003 (3.1)	−0.0002 (1.8)
Married	0.052 (1.1)	0.104 (2.5)	0.063 (1.4)	0.083 (2.0)
Agriculture	−0.700 (15.1)	−0.684 (14.7)	−0.775 (16.9)	−0.746 (16.5)
Public sector	−0.404 (8.9)	−0.254 (3.7)	−0.389 (8.6)	−0.238 (3.5)
Constant	−1.73 (24.6)	−1.61 (21.9)	−1.34 (21.9)	−1.32 (20.8)
Adjusted R^2	0.318	0.313	0.332	0.334
N	1918	1393	1918	1393

Note: t-values in parentheses with robust standard errors. Regressions (1) and (2) use continuous variable for education, while (3) and (4) use dummy variables.

Table 3 here presents the regression coefficients, in a format matching that of Lall and Sakellariou's Tables 7 and 8 (2010, 342, 343)—dummy variable coef-

4. Of the employed persons in the 2010 data set, 23 percent (767 of 3,311) are either younger than 22 years old or older than 65.

ficients are unadjusted. It is worth noting that the negative coefficient for public service work remains in the 2010 results. Lall and Sakellariou (2010, 344) hypothesized that the negative coefficient might be due to greater job security and/or less demanding work conditions. Another possibility worth considering is that people are willing to accept public sector jobs at a lower wage because they are aware that working in the bureaucracy will provide opportunities to demand bribes, favors, or 'service fees' from clients. The above data does not provide sufficient data to be able to test these different theories.

Table 4 reports the annualized education premiums based on my regressions for 2010, using the Kennedy transformation. The table reports these premiums alongside the premiums I calculated from Lall and Sakellariou's regression results (as also reported in Tables 1 and 2). Consistent with the earlier years, the results based on 2010 data show an increasing return on education, though in 2010 this is even more pronounced. Compared to the 2007 results, the education premium in 2010 is smaller for early education (primary and junior high) and larger for higher education (senior high and university). Unfortunately, the differences between the 2007 and 2010 models make it difficult to interpret these changes.

TABLE 4. Annualized education qualifications premiums, among persons employed for wages

	Men ages 22–65			Men, any age
	1997	2003/04	2007	2010
Primary	1.3%	10.4%	7.3%	5.8%
Junior high	4.6%	11.9%	10.8%	5.7%
Senior high	4.3%	11.5%	11.4%	22.8%
University	11.1%	31.2%	26.6%	40.9%
	Women ages 22–65			Women, any age
	1997	2003/04	2007	2010
Primary	0.2%	8.1%	8.8%	3.5%
Junior high	1.8%	10.1%	7.5%	6.2%
Senior high	−3.5%	13.3%	12.5%	22.9%
University	0.4%	38.6%	28.7%	34.8%

Note: Kennedy (1981) transformation of: for 1997, 2003/04, and 2007, regression coefficients as provided by Lall and Sakellariou; for 2010, regression coefficients as reported in Table 3 column (3).

The main conclusions that we can draw from the 2010 results are that the Cambodian education premium is convex (increasing returns) and that the university premium is high by international standards.

Appendix

Data and code for the results found in this paper can be downloaded **here**.

References

Griliches, Zvi. 1971. Hedonic Price Indexes for Automobiles: An Econometric Analysis of Quality Change. In *Price Indexes and Quality Change: Studies in New Methods of Measurement*, ed. Griliches, 55–87. Cambridge, Mass.: Harvard University Press.

Halvorsen, Robert, and Raymond Palmquist. 1980. The Interpretation of Dummy Variables in Semilogarithmic Equations. *American Economic Review* 70(3): 474–475.

Heckman, James J., Lance J. Lochner, and Petra E. Todd. 2006. Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond. In *Handbook of the Economics of Education*, vol. 1, eds. Eric A. Hanushek and Finis Welch, 307–458. Amsterdam: North-Holland.

Kennedy, Peter E. 1981. Estimation with Correctly Interpreted Dummy Variables in Semilogarithmic Equations. *American Economic Review* 71(4): 801.

Lall, Ashish, and Chris Sakellariou. 2010. Evolution of Education Premiums in Cambodia: 1997–2007. *Asian Economic Journal* 24(4): 333–354.

Lucas, Robert E. B. 1977. Hedonic Wage Equations and Psychic Wages in the Returns to Schooling. *American Economic Review* 67(4): 549–558.

About the Author

John Humphreys is a Visiting Lecturer at the University of Queensland, Australia, and the Executive Director of the Professional Research Institute for Management and Economics, Cambodia. His email address is john@johnhumphreys.com.au.

Go to archive of Comments section
Go to September 2015 issue