Effect of English Ability on Income in Siem Reap, Cambodia Frontline Tourism Employees

Jeffrey Stewart Morrow, PhD

Prefectural University of Kumamoto

Kumamoto, Japan

Abstract: In the past decades, researchers have discovered that general education and socioeconomic aspects like home-life, marriage, location of birth, age of arrival into a developed country (in the case of immigrants), and the like have influenced income in tourism and other industries. Research into English language education and proficiency and its role in *developing* frontline tourism income has been lacking. The purpose of this paper is to examine the effect of English ability on income through regression analysis using survey results by the author and research team in 2012 in Siem Reap (SR) tourist industry (TI). Results found positive associations between English ability level, years of English education, hours of English education, and total years of schooling although direct effect on income was low.

Keywords: English ability, income, developing country, tourism industry, Cambodia

1. Introduction

Research on the role of English ability on income began in the 1980's, and the economic effect of English on income and employment and has been investigated more recently. When examining Mexican migrants into the U.S., Sandford (2002) found that returns for those with English language proficiency and some education are higher than those with English ability and less education. Similarly, Kim (2003) found theoretical underpinnings that higher educational attainment is associated with higher incomes in a study of Chinese and Mexican male immigrants into the U.S. The role of English proficiency in South Africa, a country where English is the language of business and politics, has been researched by Casale and Posel (2010). This study focused on the ability to *read and write* English *very well* as the benchmark for proficiency. The role of English proficiency on earnings among African men between ages 25 and 65 was tested in the study. Statistical results showed that those English language speakers of high ability earned almost 55% higher incomes than non-proficient English users. Ono and Savodny (2007) studied the role of English ability in IT among immigrants to the U.S. and found lower IT usage rates, as well as incomes, by those with lower English ability. Bleakley and Chin (2003) focused on English proficiency and earnings of childhood immigrants to the U.S. The study found that a large number of immigrants who arrived in the U.S. as children were able to learn English more readily than those who arrived later. Bellante and Kogut (1998) discuss the earnings of immigrants based on the effects of language ability and working experience. Those respondents who had a reported ability of *"not well,"* earned 20% less than the base group, and

those who reported *"not at all"* earned 27% less. Morrow (2014) has found corresponding ascending values in English proficiency and income in frontline tourism staff in Siem Reap, Cambodia, suggesting an importance of English proficiency in employment and income there. DiPaolo & Tansel (2015) have discovered high economic value of foreign language skills in Turkey, a country that has international trade, high tourism arrivals, and many FDI projects. Empirical test returns to English proficiency are positive and increasing between 20 and 46%.

2. Survey Method, Dates, Location, Sample

The author conducted a tourist industry survey in 2012 in Siem Reap to examine the effect of English ability on income. Because of its proximity to Angkor Wat, Siem Reap was chosen as the survey locale because it is a typical cultural tourism destination. As such, it attracts visitors from around the world annually, and many of them are from inner circle countries and use native English or English as an International Language (EIL) in travel. During earlier travel experience there in the early part of 2000's, the author found that English communication was difficult for tour guides and realized that this could be vastly improved. In addition, Cambodia as a developing country has very little data on English education and its contribution to per capita income, which ideally lends itself to more research in the future.

2.1 Method

Manny take it for granted that English is necessary in the tourist industry overall; however, there has been very little empirical research to verify that notion. In order to more fully investigate the effect of English ability on income, the author led a research team consisting of five indigenous English-speaking students with advanced level 5 proficiency from Angkor University in Siem Reap. To conduct this research, simple stratified random surveys in a face-to-face interview format was used, and this allowed an almost perfect response rate. The sample strata were 262 tourism employees in five obvious tourist industries in 2012: souvenir shops, restaurants, guesthouses, hotels, and travel agencies. These businesses were selected because they are obvious tourism-related businesses.

The questionnaire contained questions such as gender, living situation, marital status, years of schooling, years of English education, frequency of English speaking visitors per month, and rate of English usage per month. To measure English proficiency, the author designed an assessment system ranging from 0 (no ability) to 5 (greatest ability) in order to assess English ability in a quantifiable fashion. This assessment was loosely based on the Council of Europe Framework of Reference for Languages. In this way English proficiency is easily categorized and easy to correlate and analyze numerically. In this manner, it is a very convenient way to explain English proficiency in order to coordinate it with types of jobs and employment seeking. To gather English proficiency data, the assistants were provided with hand held IC recorders to record respondents' self-introductions in English. Linguists further verified the English ability of the employees post survey to obtain English proficiency assessments. The CEFR and the author's assessment can be found in Tables 1 and 2.

A1 Breakthrough or beginner - can understand familiar everyday expressions, can introduce self

A2 Way stage or elementary - can understand and use expressions within immediate relevance

B1 Threshold or pre-intermediate - can understand main points and deal with traveling

B2 Vantage or intermediate - can understand and produce ideas on concrete and abstract topics

C1 Effective operational proficiency - can get implicit meaning, can use language flexibly

C2 Mastery or advanced - can understand everything, can express spontaneously and fluently

Table 2. CEFR Reference Levels

Table 1. CEFR Reference Levels for English.

Source: adapted from Council of Europe for Language Education, 2001.

0 No or little ability - can't communicate at all

1 Beginner - can only understand and use familiar everyday expressions with no confidence

2 High Beginner - can understand and use expressions within everyday relevance, little confidence

3 Intermediate - can understand many things, can produce but with many mistakes, less confidence

4 High Intermediate - can understand and produce ideas but with lower confidence

5 Advanced - can understand everything and produce fluently and confidently

Table 2. The Author's Variation of English Language Assessment
Source: Author, 2010.

2.2 Survey Dates and Location

The survey was conducted from March 23-25, 2012 in the following areas (Figure 1). Souvenir shops were located in Old Market only. Restaurants were located on Pub Street and Pub Street Alley, guesthouses were located on Street 2, Street 3, Oum Khun Street, and around Wat Bo Road. Travel and tour guide agencies were located on Oum Khun Street and Street 5, and Sivath Boulevard. Hotels were located on National Road 6, and Angkor Wat Road. Fortunately, regulations are not strict in Cambodia. Permission to interview was received relatively easily, and the subjects were interviewed directly at their places of work.
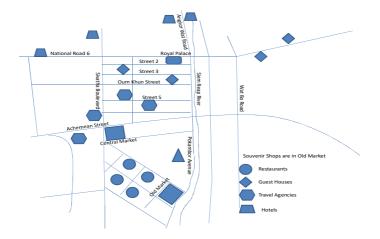
Figure 1. Map of Siem Reap Survey Areas

Source: Rendered from survey data, 2012

2.3 Sample size

The sample size was determined in two stages after obtaining preliminary data as to total number of shops in the survey bracket within Siem Reap City (Table 1). Data on total TI related working population was difficult to find at that time, so the sample size was calculated in two stages. First, an appropriate sample for the total number of locations was calculated. Then the individual shops were chosen randomly, and within those locations the participants were chosen randomly, calculated at a confidence level of 99%, and a confidence interval of 15. The strata (n=262) were among the working age population (15-65) in the SR TI in addition to some employees outside the working age population at ages 10-15. The sample participants contacted on survey day, upon entering the establishment. All participants fortunately agreed to be interviewed for this study. In order to complete this study adequately, the strata were interviewed face-to-face, and answers were written carefully by the assistants, ensuring that accurate data were collected.

Establishment	Total locations	Sample (2012)
Souvenir shops	300 (after tabulation)	60
Restaurants	120	48
Guesthouses	227 (3,000 rooms)	55
Hotels	120 (8,723 rooms)	45
Travel Agencies	142	54
Total		n=262

Table 1. Sample, 2012

Source: Department of Tourism, Siem Reap and author's calculation, 2012.

3. Results

After collecting data, the answers were coded and entered into excel, then regression analyses were employed in this analysis in order to more fully extrapolate meaningful results on the effect of ECA and income. Three separate simple linear regressions for all TI employees in all businesses put together were conducted and are found in this section. After this, the author then performed a multiple regression using other English related variables from the study. For the simple regressions, income is used as the dependent variable, while English level, total years of English education, and total years of schooling as the independent variables. These were chosen as the variables because descriptive statistics showed they were the most impactful for the quality of the work situation in SR TI and the most directly related to income. In this manner, the author could verify the actual effect of those variables that were directly income-related, and have no indirect or secondary relationship.

The regression results of income as dependent variable and English level (Eng lvl) as independent will be offered now (Table 2). *Eng lvl* refers to the English proficiency level as numeric value. The regression equation is y=61.78+14.42x. As can be seen in the resulting outcomes, P values for both the constant and English level variable are significant. R, or the correlation coefficient, was .34, which is a decent correlation and can be considered significant. R^2 value on the other hand is significant (.11), but the value indicates that only 11% of income is accounted for by English level. The regression as a whole is significant at <.001 level and denotes that the variables did not occur by chance. Therefore, the null hypothesis can be rejected. It can also be discerned from the regression that the y-intercept and coefficients are significant (<.001). Coefficients show the effect of change of independent variable on the dependent variable. In this data, the coefficients indicate a positive effect relationship. Based on these results, it is quite obvious that English level and income have a positive relationship in SR TI, although the direct effect of *Eng lvl* on income is low.

	Mean	SD (stnd error)	t-Stat	P	Lower	Upper
Income	102	49.73(7.56)	8.17	.001*	46.9	76.67
Eng Lvl	2.78	1.71(2.50)	5.76	.001*	9.5	19.35

R 0.33, $R^2$0.11, *P<.001
Table 2. Income and English level
Source: Calculated from Survey Data, 2012

The regression results of income as dependent variable and years of English education (yrs Eng edu) as independent variables are now given (Table 3). *Yrs Eng edu* refers to years of English education in number of years. The computed regression line for income and years of

English education is y=90.61+4.07x. Here, income is the constant and independent is years of English education. The correlation coefficient is .18, and can be considered significant, which satisfies the requirement that the results did not occur by chance. R^2 value in this regression is positive; however, again we can only determine that 3% of income is accounted for by years of English education. Values are significant coefficients for each variable indicating that our y-intercept and independent variables are valid. This regression as a whole is significant at <.001 as well and the null hypothesis can be rejected. The P values for the intercept and the variable, *Yrs Eng Edu* can be considered significant at <.001, and therefore show validity. Again, this data show that there is a positive relationship between years of English education income, although direct effect of *Yrs Eng edu* on income is low.

	Mean	SD (stnd err)	t-Stat	P	Lower	Upper
Income102	49.73(4.48)		18.71	.001*	81.08	100.14
Yrs Engedu	2.8	2.24(1.34)	3.02	.002**	1.41	6.73

R 0.18, $R^2$0.03, *P<.001, **P<.05

Table 3. Income and years of English education

Source: Calculated from Survey Data, 2012

The regression results of income as the dependent variable and total years of schooling (Ttl yrs scho) as independent are now offered. *Ttl yrs scho* refers to the total years of schooling in number of years. The regression equation is y=55.43+3.87x. The correlation coefficient, R, is .24 and the R^2 is 0.05, again indicating that only 5% of income is accounted for by total years of schooling. The author hoped a higher predictor for total years of schooling. Values for the coefficients are significant and P values for each coefficient is significant at <.001. The whole regression is significant at <.001 as well, and the null hypothesis can be rejected. It can be said that years of schooling have a positive relationship with income in this data, although the direct effect of *Ttl yrs scho* on income is low.

	Mean	SD (stnd err)	t-Stat	P	Lower	Upper
Income102	49.73(12.01)	4.61		.001*	31.77	79.09
Ttl scho yrs	12.02	3.09(0.96)	4.00	.001*	1.96	5.78

R 0.24, $R^2$0.05, *P<.001

Table 4. Income and total school years

Source: Calculated from Survey Data, 2012

In order to more fully examine the effect of English background on income empirically, the author performed a multiple regression for all TI employees together. In this regression, income

was the constant and the independent variables were as follows: *Edu lvl* - a variable where 1 equals primary school only, 2 is equal to secondary school, and 3 equals high school graduate, and 4 is equal to university graduate; *Ttl yrs scho* refers to the total years of schooling in number of years; *Yrs Eng edu* refers to years of English education also in number of years; *Hrs Eng edu* refers to the hours of English education in school in number of hours; *Freq Eng sp* refers to the frequency of English speaker visits in days per week; *Eng lvl* refers to the English proficiency level as a numeric value; *Exp wEng* refers to the expenditure of monthly English study in US \$; and *Eng day/mo* refers to English usage in days per month. The multiple regression equation is: $y=25.87+10.40x_1+-0.57x_2+1.99x_3+-2.68x_4+10.40x_5+13.21x_6+-0.18x_7+0.40x_8$

Multiple regression results show quite significant findings with respect to education and income in SR TI. The entire multiple linear-regression is significant ($<.001$), which denotes the regression values did not occur by chance, and the null hypothesis can be rejected. The correlation coefficient is .40 and the R^2 value is higher than the previous two regressions at .14, although this regression indicates that the independent variables account for only 14% of income. Some independent variable results in the multiple regression indicate *insignificant* P values: *Edu Lvl* (0.34), *Ttl yrs scho* (0.85), *Yrs Eng edu* (0.18), *Freq Eng Sp* (0.22), *Exp wEng* (0.22), *Eng day/mo* (0.56). Significant coefficients are *Hrs Eng Edu* ($P<0.001$) and *Eng Lvl* ($P<0.001$). These factors being the case, we can clearly see that two independent variables (Hrs Eng Edu, Eng Lvl) have a positive relationship on income, although again, *direct* effect on income by *Hrs Eng edu* and *Eng Lvl* is low.

	Mean	SD (stnd err)	t-Stat		P	Lower	Upper	
Income	102	49.73(31.17)	0.83		0.40	-35.51	87.26	
Edu lvl	2.86	0.91(10.94)		0.95	0.34	-11.15	31.96	
Ttl scho yrs	12.02	3.19(0.95)	-0.17		0.85	-6.86	5.71	
Yrs Eng edu	2.8	2.24(1.48)		1.33		0.18	-0.93	4.91
Hrs Eng edu	4.19	3.24(1.01)		-2.6		0.008*	-4.68	-0.67
Freq Eng sp	1.34	0.74(8.49)		1.22		0.22	-6.33	27.13
Eng lvl	2.78	1.17(3.01)		4.38	.001*	7.28	19.15	
Exp wEng	10.86	19.53(0.15)		-1.21		0.22	-0.49	0.11
Eng day/mo	25.42	9.21(0.68)		0.58		0.56	-0.95	1.75

R 0.24, R^2 0.05, *P<.001

Table 5. Multiple Regression Results
Source: Calculated from Survey Data, 2012

8. Discussion

During the regression, both useful and inexplicable things occurred. Firstly, the analyses resulted in significant values in both simple and multiple regressions in all cases. The R^2 values are low in some cases, which in some cases could mean that the data does not fit the model, and in some cases resulted in the variable explaining only a small portion of the dependent variable. However, very little research has been completed in English ability in tourism in developing countries, so there is relatively little literature with which to compare this study.

Regression results for English level in the first simple regression were significant ($P<.001$) and imply that English ability was important for income. In the English ability regression the R^2 is low which shows our data does not fit the model 100%, but one must keep in mind that this research is dealing directly with people, not science, and many factors influenced the survey situation. Some of these factors include whether or not the subject answered correctly, the interviewing of the assistants, and interpretation of the questions by the respondents. The procedure of face to face interviews did not help this matter; however, the author feels that face to face interviews that allow the participate to answer questions him or herself, relieving the interview of interviewer bias, are the most accurate way to proceed. In fact, many mail-in questionnaires have lower participation rates, and therefore, lower results. These results indicate that English proficiency influences income in our data. Further evidence is that the regression as a whole denotes that the variables did not occur by chance, is significant ($<.001$), and the null hypothesis can be rejected. P values show that the y-intercept and coefficients are valid indicating robustness in the first regression. Coefficients show the effect of change of independent variable on the dependent variable. In our data, the coefficients are positive, indicating a positive effect relationship. Based on these results, it is quite obvious that English level has a positive effect on income in TI in SR. *Eng lvl* is important in getting many jobs in SR, and even though the direct effect on income is low, the results are still important.

The next simple regression using income as the dependent variable and total years of schooling as the independent show that the R^2 is .05 is still positive and important to give proof to the findings. P values are $< .001$, showing validity in the y and independent variable. Education is extremely important in developing countries, it can be surmised that this level is significant and has a positive effect on income and employment. The direct effect on income is low, but this may indicate that many young people have little chance to obtain a sufficient education.

In the simple regression of income as the dependent variable and English education as the independent variable, the R^2 is very low, and this may indicate that the regression model is not a good fit. However, P value is highly significant for the entire regression, and it is also significant for the intercept (income) and the years of English education variable. The reason for the low R^2 is that data collection in developing countries is difficult due to many inconsistencies

that could occur in the data collection situation, such as difficulty in answering questions, inconclusive answers, inaccurate local data, miscommunication, and the like. Here again, the regression shows significant F at < .001 and again the null hypothesis can be rejected. Again we can say with certainty that years of schooling has a positive influence on income in our data, although even perhaps less so that English proficiency and education. This regression results point to an interesting situation in a developing country where English proficiency, English education, and total years of schooling have great impact on incomes, employment, and livelihoods. Without English proficiency and education, young Cambodians cannot obtain adequate salaries for life. Even though the salaries are still far from being comparable to Western countries, there is still important impact on income for those with English ability. Empirical results by both Becker and Mincer show that schooling has been a major positive role in income and jobs in peoples' lives. This fact is difficult to prove in Cambodia due to little data, uncertain political situations, and conditions that are not standardized. This is certainly the case in our survey situation; however, these results prove that human capital attainments in English proficiency, English education, and general schooling, all have positive effects on the income situation in Cambodia, especially in tourism.

In the multiple regression analysis, years of English education values are also significant, although R^2 displays a less than accurate fit of the model. Upon examining the R^2 value, we find that around 16% of income is explained by the independent variables. One may think this is a low value, but it is not necessarily cause for concern. The lowness can be explained by the fact that in Cambodia, there is much variability in all aspects of life; the TI has no minimum wage, standardized salaries, nor set salary. It goes without saying that some people have these items, but not all. Even knowing this, however, the R^2 value is acceptable for the situation in Cambodia as all developing countries as most have little data and most correlations, especially to education, return negative values. In addition, this study examined human data, not scientific. As such, many things must be taken into account when interpreting the results. Multiple regression results showed that there is some effect of our independent variables relating to English background and income. Significant P values were found in coefficients *Hrs Eng edu* (P<0.00) and *Eng lvl* (P<0.00); however, the P values are not significant in some other coefficients: *Edu lvl* (0.34), *Ttl yrs scho* (0.85), *Yrs Eng edu* (0.18), *Freq Eng sp* (0.22), *Exp wEng* (0.22), *Eng day/mo* (0.56). T-stats were significant in the same variables as well. Eng lvl values showed significant P values (<.001), and significant F was found for the whole regression. From these findings we can stipulate that *Edu lvl, yrs Eng edu, freq Eng sp, Eng lvl,* and *Eng days/mo* all had influence on the dependent variable. Even though the R^2 values were not high, it still has positivity and is slightly robust for the TI situation in SR. There are many factors which influence the survey situation in developing countries, and it should not be taken that our regression results are meaningless. The end product here is that regressions are somewhat useful in giving an association of the effect of ECA on income; however, it does not tell much more.

For example, in this very social-survey situation, perhaps it is more useful to use descriptive and correlation for analysis.

10. Conclusion

In conclusion, this empirical evidence shows that larger returns to investments in English education and proficiency can be obtained within the tourist industry in Siem Reap. The results are not outstanding but show some relationship with the power of English proficiency and education. The returns to English ability are higher in TI than in agriculture, garment, retail, or manufacturing. The study confirmed that hours and years of English study in school, in addition to money spent on learning English monthly, and higher English proficiency are paramount to receiving better jobs, aspects that are directly related to income in tourism. Even those elements indirectly related to income, such as frequency of English speaking customers and usage of English per month, are beneficial to the young labor force as more young people enter the TI job market. This survey was somewhat a pioneer study as none like it has been done in the past. Although the sample size small and the survey data collection situation difficult, the research team was able to gather some useful socioeonomic information from the present tourist industry situation in Siem Reap. In addition to designing 2 questionnaires and collecting socio-economic data, the author made several inroads into the discovery of the impact of English ability on income and employment generation in a developing country. The survey experience led to more understanding of the socio-economic situation and how to foster income and employment growth in a typical cultural tourist destination in a developing country.

Regression results showed that there is some influence among the independent variables relating the effect of English proficiency and education on income, as well as among total years of schooling on income, even though the direct effect on income was low in all cases. In the simple regression of income and English level, the R^2 values are highest among the 3 in regression 1. Therefore, this indicates that English proficiency influences income in our data. Further evidence is that F, which denotes that the variables did not occur by chance, is highly significant at the <0.001 level. P values show that the y-intercepts and coefficients are both valid and show the influence of the independent variable on the dependent variable. Coefficients, which illustrate the effect of change of independent variable on the dependent variable, are positive in this data, indicating a positive influence from the independent variable. Based on these results, it is quite obvious that English level had at least an indirect, though positive, effect on income in TI in SR. In the regression of English education and income, the R^2 is not as high as regression 1, but it is still positive and gives some weight, albeit small, to our findings. Again, P values are quite significant as well, showing validity in the y and independent variable, as are the coefficients. This indicates a strong influence of years of English education on income. In the total years of schooling and income regression, although not as robust as regression 1, shows a slight positive relationship of schooling on income, although a small effect. Here again, the whole regression is significant at <0.001 and again the null hypothesis can be rejected. P values

are significant as well and show validity in both the dependent and independent variables. Again we can say with certainty that years of schooling has a positive influence on income in our data, although even perhaps less so than English proficiency and education.

In the multiple regression, we found positive coefficients in *edu lvl, yrs Eng edu, freq Eng sp, Eng lvl,* and *Eng days/mo*. P values showed significance in all independent variables, and significance F was found for the whole regression. From these findings we can verify that *edu lvl, yrs Eng edu, freq Eng sp, Eng lvl,* and *Eng days/mo* all had a positive effect on income. Even though the R^2 values were not high, it still has positivity for the TI situation in SR. There are many factors which influence the survey situation in developing countries, and it should not be taken to mean the regression results are meaningless. From these findings it is quite clear that English education and ability can generate help better income, and furthermore in employment prospects. This can be seen in the regression analysis, in the income distribution, in the employment generation, and in the attainments of English proficiency in the five businesses under study in SR. Of course, we cannot say that English background alone leads to higher incomes in *every* case in *all* countries, but clearly there is a positive influence of English on income and employment at least in TI in SR from these findings.

From this study, some factors of general and English background education, the socio-economic situation, income, family life, and living conditions of those working in tourism in a developing country were on the next step to authentification. The goal was to examine correlations of income with English proficiency and with other variables to examine how to create better income and employment situations in its role in employment, and even though small direct effects on income from English related variables were found, positive relationships were exhibited in this outcome. Other researchers should more fully examine the connection between English proficiency and employment in the future. This will enable researchers to understand the true impact of language proficiency and communication on the lives of those in developing countries.

References

Author. (2014). The Role of Human Capital Attainment for Generating Employment

and Income: The Case of Tourist Industry in Siem Reap, Cambodia. *Kumamoto Gakuen University Journal of Economics,* 20 (1-4), 103-159.

Bellante, D. & Kogut, C. (1998). Language Ability, US Labor Market Experience and the Earnings of Immigrants. *International Journal of Manpower,* 19:5, 319-330.

Bleakley, H. & Chin, A. (2004). Language Skills and Earnings: Evidence from

Childhood Immigrants. *The Review of Economics and Statistics,* 86(2), 481-496. doi: 10.1162/003465304323031067

Casale & Posel. (2011). English Language Proficiency and Earnings in a Developing Country. *Journal of Socioeconomics*. (40)4, 385-393.

DiPaolo, A. & Tansel, A. (2015). Returns to Foreign Language Skills in a Developing Country: The Case of Turkey. *Journal of Development Studies*, DOI:10.1080/00220388.2015.1019482

Ono, H. & Zavodny, M. (2008). Immigrants, English Ability and the Digital Divide. *Social Forces,* 86:4, 1455-1479.