Economics & Management Publications   Faculty and Staff Scholarship

5-25-2021

What Factors in Nepal Account for the Rural-Urban Discrepancy in Human Capital? Evidence from Household Survey Data

Satis Devkota
University of Minnesota - Morris, sdevkota@morris.umn.edu

Shankar Ghimire
Western Illinois University

Mukti Upadhyay
Eastern Illinois University

Recommended Citation

Article

What Factors in Nepal Account for the Rural–Urban Discrepancy in Human Capital? Evidence from Household Survey Data

Satis Devkota [1], Shankar Ghimire [2,*] and Mukti Upadhyay [3]

1 Social Science Division, University of Minnesota Morris, Morris, MN 56267, USA; sdevkota@morris.umn.edu
2 School of Accounting, Finance, Economics, and Decision Sciences, Western Illinois University, Macomb, IL 61455, USA
3 Department of Economics, Eastern Illinois University, Charleston, IL 61920, USA; mpupadhyay@eiu.edu
* Correspondence: sp-ghimire@wiu.edu; Tel.: +1-309-298-1153

Abstract: We analyze the factors that determine human capital formation in the rural and urban sectors of Nepal and decompose the intersectoral difference into variables underlying supply and demand for human capital. In particular, we examine the role of access to primary and secondary schools as well as the socioeconomic, demographic, and geographic characteristics of households. Our results are based on Nepal Living Standards Survey data for 2004 and 2011. We find that access to schooling has a significant impact on the level of human capital, especially in rural areas. Our Blinder–Oaxaca decomposition attributes a large portion of the rural–urban gap to socioeconomic and demographic variables. Yet, the results reinforce our claim that an improvement in schooling access and road infrastructure is also necessary, particularly in the vast rural sector of Nepal, if human capital development is to provide a greater contribution to national welfare.

Keywords: human capital index; rural–urban difference; Blinder–Oaxaca decomposition; Nepal Living Standards Surveys

1. Introduction

This paper identifies the determinants of human capital formation in Nepal and calculates and decomposes the difference in human capital into supply and demand side factors. In particular, it examines the role of supply side factors such as access to primary and secondary schools and demand side factors such as the socioeconomic, demographic, and geographic characteristics of households. In doing so, we use the Nepal Living Standards Survey Data for the years 2004 and 2011 and construct the level of human capital for the pooled sample. Then, we estimate the effect of the access (measured in terms of the travel time to the nearest school) and other covariates on human capital formation in the urban and rural subsamples. This allows us to estimate the difference in the predicted value of human capital between the sectors and analyze the extent to which this difference is explained by differences in the supply and demand factors listed above.

The rural–urban difference in human capital is a topic of high relevance for economic policymaking, yet it remains under-researched in empirical economics. The topic is even more important for developing countries such as Nepal, where the majority of the population (80%) still lives in rural areas lacking in physical infrastructure in health, education, electric power, transportation, drinking water, and sanitation. According to government data, public investment in infrastructure has fluctuated over the years (6.6% of GDP in 2007/08, 3.4% in 2010/11, 10.1% in 2016/17). Around the same time, government expenditure on education has increased from 3.5% of GDP in 2007 to 5.5% in 2017, which is among the highest in South Asia (World Bank 2017). However, the majority of the government expenditure predictably goes to the urban areas, which is likely to increase the average

human capital among urban population compared to the rural population. In our data, the overall urban–rural difference in human capital is 0.3065. This is a significant difference when placed in the context of an overwhelmingly large rural population.

To what extent the supply factors and demand factors are responsible for the difference in human capital is not clear a priori. For a balanced growth between urban and rural economies, however, the human capital discrepancy has clear policy implications (Chattopadhyay 2012, 2014). If so, could all of this discrepancy be explained by unfavorable socioeconomic and demographic characteristics in villages, or does the availability of schools, and in particular, the time it takes children to reach the nearest school (the access), matter as well? The novelty of this paper, thus, lies in the estimation of the effect of access to schooling as well as the role of socioeconomic, demographic, and geographic factors on human capital formation. We estimate our model for urban and rural subsamples separately and calculate the predicted difference in human capital. This difference is then decomposed using the Blinder–Oaxaca method (Blinder 1973; Oaxaca 1973) into parts attributable to various explanatory factors. We also perform a robustness check using the instrumental variable regression technique for parameter estimation as well as decomposition of the rural–urban difference in human capital.

A key contribution of this paper is to test how the socioeconomic, demographic, and geographic conditions are also responsible for human capital formation. We conduct this analysis accounting for both the supply and demand side factors reflecting their relative roles. Our empirical strategy also includes instrumental variable regression because of the endogeneity of income, a potent demand factor. With regard to our key variable of interest, we find that access to schooling significantly explains the difference in rural and urban human capital formation. Yet, much of the difference between the two sectors is indeed accounted for by the household's socioeconomic and demographic factors. We find that improvements in the demand factors considerably reduce the rural–urban gap. Based on our knowledge of the existing literature on Nepal, a decomposition of the rural–urban difference in human capital has not been studied for the country yet. We hope to contribute to the literature in a way that directly leads to significant policy implications.

The remainder of the paper is organized as follows: Section 2 reviews the literature that examines the role of access to education and other demand factors on human capital formation; Section 3 explains the theoretical framework in calculating the human capital index; Section 4 presents the data and the empirical methodology used in the analysis including the Blinder–Oaxaca decomposition; Section 5 presents and discusses the empirical results; and Section 6 concludes.

2. Literature Review

Human capital formation is an important topic in economic growth and wellbeing. Existing studies on the topic identify demand side factors such as income and other household characteristics as affecting human capital formation (Bansak and Chezum 2009; De and Ratha 2012). The limited studies that analyze the supply side factors focus on the impact of access on educational outcomes of children, such as school enrollment rates, test scores, and grade attainment (Handa 2002; Mejía and St-Pierre 2008). We extend these analyses further to identify the relative roles of the supply side and demand side factors on human capital formation.

On the supply side, existing research shows a positive impact of rural road construction on household income and income inequality by providing better access to markets (Charlery et al. 2015). Along the same line, Jacoby (2000) shows a positive impact of road access to markets in rural Nepali communities. We extend this idea to analyze if increased access to educational institutions would contribute to one's human capital. As the increasing access to the road is a pro-poor intervention, it increases access to schooling and health care services for the rural poor. Thus, these results help us to frame our hypothesis that a difference in rural–urban infrastructure as measured by the access to schooling explains at least part of the rural–urban differences in human capital.

On a related note, Sen (1985) argues that access to education and other areas of infrastructure opens doors to multiple economic opportunities and makes people capable. Expanding optionality or choice enables people to lead a more prosperous life by improving the three essential factors that contribute to people's wellbeing: knowledge, long and healthy life, and access to resources needed for a decent standard of living. Existing research supports such claims by showing that socioeconomic, demographic, and geographic factors (mainly age, health, and income) equally influence children's health and educational achievement (Akee et al. 2018; Currie and Almond 2011; Currie 2009; Cunha et al. 2006). Thus, the literature suggests a balanced role for both the supply and demand side factors to enhance human capital accumulation. For example, Banerjee and Duflo (2011) present interesting stories of schools being available for free in most developing countries, for at least the primary level. They report that most children are enrolled in school, and yet, in the various surveys they have conducted around the world, child absentee rates are high, ranging from 14 to 50 percent. They claim that the failure of schools in developing countries to attract children cannot be explained by the problems of access only. The parents do not seem to be able or willing to make their children go to school. That is, the demand side factors are also equally important.

The importance of access to infrastructure on overall economic development is well established in the policy arena across the globe. Multilateral development partners such as the World Bank and the World Trade Organization (WTO) have argued for prioritizing the infrastructure sector for overall development strategy. The 1994 World Development Report had found infrastructure development to be a most effective instrument of direct help to the poor. Among many other studies, Calderón and Chong (2004) find that an increase in both the quantity and quality of infrastructure reduces income inequality in a cross-section as well as a panel of countries, with the quantitative link stronger in developing nations. Infrastructure helps reduce poverty, increases school enrollment rates, decreases school dropout rates, and enhances the utilization of health care services. Each of these changes will increase human capital formation.

Again, at the World Bank, Leipziger et al. (2003) show that apart from traditional variables (income, assets, education, and direct health interventions), better access to facilities significantly improves child *health outcomes*, a major determinant of human capital. Another line of research analyzes the impact of infrastructure directly on *human development and quality of life*. For instance, Sapkota (2014) presents a cross-country analysis of the impact of several infrastructure measures (access to electricity, access to clean drinking water, and road density) on the Human Development Index (HDI) and its components individually.

To focus our attention on Nepal, the infrastructure state of the country has many dimensions. Nepal ranks near the bottom (168th) of a list of 171 countries in the Logistics Performance Index measuring the quality of trade and transport-related infrastructure (World Bank 2017). According to a ranking by the World Economic Forum (2015), Nepal ranks 132nd out of 140 countries in infrastructure development based on the joint index of transportation, electricity, and telephony networks. This lines up well with the country's record of having the slowest economic growth among South Asian countries, together with a high poverty headcount as well as income inequality in the region (Gilbert and Banik 2012). In fact, according to these authors, a reduction in transport margins, which would signify infrastructure development, would benefit Nepal the most among the South Asian Sub-regional Economic Cooperation (SASEC) countries. In quantitative terms, a 20% reduction in transportation and processing time would increase household welfare by 12% of GDP and lead to a significant improvement in human development.

As is evident from above, the extant literature has considered the impact of access to infrastructure on individual components of human development such as health, education, and income. Our research focuses on the impact of access to education infrastructure, which is likely to have a direct effect on human capital formation as well as an indirect effect via socioeconomic factors. Hence, we first identify the rural–urban difference in human capital and analyze if that difference is explained by the difference in access to schools and other

demand side factors such as socioeconomic, demographic, and geographic conditions that might, for instance, help increase school enrollment, reduce school dropouts, improve student performance, etc. The following sections elaborate this further.

3. Theoretical Framework

We begin by constructing a human capital index (HCI). Three popular approaches in this part of the literature generate outcome-based, cost-based, and income-based measures of human capital. Barro and Lee (1993) adopted the outcome-based approach to use school enrollment rates as a proxy for human capital. Similarly, Nehru et al. (1993) used a related measure of the accumulated years of schooling in the employable age for human capital. Because a student's effectiveness can only be recognized when they participate in the labor market, school enrollment cannot be the best proxy for human capital.

The cost approach toward human capital measurement has been followed by, among others, Kendrick (1994), and Jorgenson and Fraumeni (1989). Here, researchers distinguish the investment and consumption aspects of education, in which investment builds human capital and consumption does not. However, it is difficult to precisely set a boundary between investment and consumption for appropriately assigning the cost of human capital. This approach is based on the returns that an individual obtains from the labor market on her educational investment even though the "human-unrelated factors" can also influence an individual's income. The cost approach, therefore, fails to accurately measure human capital.

The third (i.e., the income) approach takes the expected income stream of an individual to calculate the present discounted value of the incomes as the aggregative measure of human capital. Graham and Webb (1979) pioneered this approach to estimate human capital for U.S. males, 14 to 75 years of age, by analyzing 1970 census data. Their interesting study estimated the present value after discounting (at 7.5% per year) the expected lifetime earnings and also derived the implied rate of depreciation of human capital of an individual over time. Obviously, a difference in the estimation method can lead to significant differences in the estimates of human capital. Graham and Webb (1979) find that their estimates of human capital were enormously greater than those based on the investment cost approach taken by Kendrick (1994). The two stocks of human capital could only be reconciled by applying a discount rate as high as 20 percent to the earnings data in Graham and Webb.

In this study, we take the income approach to estimating human capital which starts from the rates of return to schooling that have now become standard in the education and growth literature. One way human capital index (HCI) can be calculated is by using the equation $h_i = e^{\phi(s_i)}$, where h_i is the stock of human capital for individual i, and s_i is their completed years of education (Woessmann 2003). Since $\phi(s_i) = r_i s_i$, where r_i is the return to education for i, the level of human capital for individual increases with the level of education augmented by the corresponding return to education. Thus, h_i can be expressed as: $h_i = e^{r_i s_i}$.

However, a single value for the return to education for all individuals is unrealistic. Assuming the returns to education are subject to diminishing returns, as shown by Psacharopoulos (1994) in his global study, the new approach measures human capital by separating the levels of schooling and the corresponding returns (UNDP (United Nations Development Program) 2015). In Equation (1), therefore, the marginal (annual) returns to higher levels of education should be less than returns to lower levels ($r_3 < r_2 < r_1$):

$$h_i = (r_1)^{s_{1i}} \times (r_2)^{s_{2i}} \times (r_3)^{s_{3i}} \tag{1}$$

where r_1, r_2, and r_3 are returns to education for the first four years of schooling (grades 1 through 4), next four years (grades 5 through 8), and the next four years (grades 9 through 12). These annual returns are $r_1 = 13.4\%$, $r_2 = 10.1\%$, and $r_3 = 6.8\%$, obtained from Weil (2016, p. 162) who relies on estimates made by Psacharopoulos (1994). The completed years of schooling that correspond to these returns are represented by s_{1i}, s_{2i}, and s_{3i}. Accordingly, Equation (1) yields our measure of human capital.

4. Empirical Framework

4.1. Data

We use data from the Nepal Living Standards Survey (NLSS) collected by the Central Bureau of Statistics (CBS), Nepal. The first round of the survey (NLSS-I) was carried out in 1995/96, the second (NLSS-II) in 2003/04, and the third (NLSS-III) in 2010/11. These surveys have followed the Living Standards Measurement Survey (LSMS) methodology advanced and promoted by the World Bank.

We use NLSS-II and NLSS-III to construct a pooled cross-sectional dataset. Data were collected using stratified random sampling by dividing the country into several strata (six in NLSS-II, and fourteen in NLSS-III). Each stratum was then divided into primary sampling units (PSUs) in all of the 75 districts in Nepal. The PSUs are wards representing the lowest administrative units in each district. A two-stage stratified sampling method was used to select the sample households. In the first stage, 326 and 499 PSUs were randomly selected for the NLSS-II and the NLSS-III, respectively. In the second stage, 12 households from each selected PSU were randomly chosen for the interview. Thus, the final sample consists of 3912 households from NLSS-II and 5988 households from NLSS-III. Our pooled cross-sectional data contain information on many dimensions of individual and household wellbeing. These include income, expenditure, employment status, educational attainment, child and maternal health status, and child immunization status. We also have information on demographic factors and the geographic location of each household.

4.2. Econometric Estimation

We estimate human capital using a regression model to determine the impact of several variables that emerge from our review of the literature as well as conditions prevailing in Nepal. Our dependent variable is the human capital that we calculated using rates of return and years of schooling, as previously noted. For each household, we consider proximity to school measured in terms of the travel time (access) as one of the explanatory variables and the variable of interest in this paper. This access variable is a supply side factor that we expect to significantly affect human capital accumulation. Following the existing literature, an individual's income, occupation type of the household head, marital status, gender, and status of health are expected to be the main socioeconomic and demographic factors affecting human capital. These factors are control variables and are listed in Equation (2) below. Administrative regions and geographic locations of the households could also affect an individual's decision regarding the acquisition of human capital and will serve as further controls in our model. Equation (2) given below captures the essential elements of our estimation strategy:

$$h_i = \beta_0 + \beta_1 acc_i + \sum_{j=2}^{k} \beta_j \mathbf{x}_{ji} + \varepsilon_i \tag{2}$$

where acc_i is the access to education measured in terms of the travel time to school and $\mathbf{x}_{ji}$ is the set of socioeconomic, demographic, and geographic location factors discussed before.

We estimate the level of human capital separately for the rural and urban subsamples. Any difference between human capital for rural areas, averaged over all rural individuals, and human capital for urban areas, averaged over all urban individuals, can be attributed to the intersectoral differences between covariate means and between covariate parameters. The fraction of the rural–urban difference in human capital to be accounted for by the difference in each of the covariates can then be determined by a decomposition technique. We use the Blinder–Oaxaca decomposition procedure for this purpose (Blinder 1973; Oaxaca 1973). The decomposition is elaborated in Equation (3):

$$h^r - h^u = \left[(acc^r - acc^u)\beta_1^r + \sum_{j=2}^{k} \left(\mathbf{x}_j^r - \mathbf{x}_j^u \right)\beta_j^r \right] + $$
$$\left[(\beta_1^r - \beta_1^u)acc^u + \sum_{j=2}^{k} \left(\boldsymbol{\beta}_j^r - \boldsymbol{\beta}_j^u \right)\mathbf{x}_j^u \right] + (\beta_0^r - \beta_0^u) \tag{3}$$

where h^r and h^u are the estimated values of human capital for the rural and urban subsamples, respectively, $\mathbf{x}_j^r$ and $\mathbf{x}_j^u$ are the vectors of sectoral means, $\boldsymbol{\beta}_j^r$ and $\boldsymbol{\beta}_j^u$ are the vectors of sectoral regression coefficients, acc^r and acc^u are the two means of the access-to-school variable, and β_1^r and β_1^u are the two regression coefficients of the access variables, respectively.

The right-hand side of Equation (3)[1] has two components after we simplify it by suppressing two other items (more below). The first part is accounted for by the actual *difference in variables* as shown by the difference in mean values for rural and urban sectors—a part that is explained by the differences in access and the group differences in socioeconomic, demographic and geographic location factors. The second part of the difference between rural and urban outcomes is attributable to the difference in variable coefficients (weighted by the respective variable means). Finally, the remaining terms are the difference between the rural and urban intercepts, and an interaction term that implies that the variable differences and coefficient differences can occur at the same time (Jann 2008). Our results section provides further details on the decomposition of the rural–urban difference in human capital.

The sample data permit the construction of two of our variables. *Access to school* reflects the average time taken to reach the nearest primary or secondary school. We construct our human capital index (HCI) using Equation (1) based on established research (Weil 2016) on returns to education. These returns are 13.4% for the first four years, 10.1% for the next four years (grades 5–8), and 6.8% for schooling beyond the eighth grade. Since NLSS gathers information at the individual level as well as the household level for the rural and urban civilian population, we are able to construct HCI for each individual in each household of rural and urban sectors.

Next, we estimate separate regressions for rural and urban populations to obtain predicted values of HCI for these broad subsamples. We calculate the difference in predicted values for the two sectors and test if the intersectoral difference in human capital is significant. What accounts for the rural–urban difference in the realized human capital? The answer is provided by the popular Blinder–Oaxaca decomposition (OBD), which we use to analyze the intergroup differences in the response variable resulting from differences in the model explanatory variables (Azam 2010; Gradin 2015, among others). This will help to understand if increased access to schooling on the supply side or various demand side factors can contribute to reducing the rural–urban gap in human capital development.

The approach we have taken toward the estimation of our regression models and decomposition of our results leads to accurate outcomes if the regression equations are properly identified. One of the explanatory variables in our human capital model (Equation (2)) is household income, which exerts a positive and significant effect on human capital formation. However, income is likely a function of the wealth of the household, the intelligence of its working members, management skills of the household head, and other factors, including human capital itself. Because of several relevant variables effectively sitting in the error term in the regression, the income variable becomes endogenous. To address the problem, we use the number of labor hours available to the household as an instrument for household income. The labor resource available to a household is highly correlated with the household income but not directly correlated with human capital formation, which makes labor hours a good instrument (IV) for household income. We use an IV regression and decompose the IV-based rural–urban difference in human capital to supplement our OLS findings and to check their robustness. To the best of our knowledge, such research on human capital in Nepal has not been performed to date.

5. Results and Discussion

Before presenting our results, it is important to note that our analysis is based on estimated human capital for the labor force ranging in age between 25 and 60. This is the normal procedure in studies of the human capital attained by the adult labor force. Table 1 shows our first result. The level of human capital is greater for the urban sample (2.5991) than for the rural (2.2926). This unconditional difference of 0.3065 is statistically significant at the 1% level. Other variables used in this study are listed and defined in Table 1.

Table 1. Descriptive statistics.

Variable Name	Definition of Variable	Pooled	Urban	Rural
h_capital	Human capital stock of individuals age 20–60	2.4111	2.5991	2.2926 ***
access_sch	Access: average time to primary and secondary schools	2.0879	0.2288	2.4362 ***
lninc	Log of income	11.6552	12.4302	11.5549 ***
occ_farmer	Occupation of the household head == Agriculture	0.5295	0.2149	0.5644 ***
occ_labor	Occupation of the household head == Labor	0.2309	0.3269	0.2199 ***
age	Age	40.0232	38.2899	40.2711 ***
age^2	Age × Age	1708.787	1564.643	1729.75 ***
sex	Sex	0.4604	0.5098	0.4696 ***
married	Married	2.8486	2.6919	2.8659 ***
excellent_SAH	Excellent Health	0.5241	0.5766	0.5183
good_SAH	Good Health	0.4488	0.4125	0.4532
privileged_caste	Privileged Caste	0.4436	0.7448	0.4216 ***
indigenous_caste	Indigenous Caste	0.3207	0.1053	0.3606 ***
central	Central Region	0.3426	0.6699	0.3245 ***
western	Western Region	0.1884	0.0984	0.2137 ***
mid_western	Mid-Western Region	0.1149	0.0411	0.1329 ***
far_western	Far Western Region	0.0717	0.0249	0.0829 ***
hills	Hill Belt	0.5179	0.7032	0.5309
mountain	Mountain Belt	0.0955	0.0171	0.1126
time	Year 2004 = 1, Year 2011 = 2	1.5794	1.6547	1.5667

Source: Calculated by authors; Standard Errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

5.1. Regression Results

The estimated coefficients and significance of results for the pooled model and for rural and urban subsamples are reported in Table 2. The R^2 for the pooled, urban and rural samples are 0.3069, 0.3060, and 0.3092, respectively.

Table 2. Regression outcomes—pooled, urban, and rural samples.

	OLS-Estimates			IV-Estimates		
Variables	Pooled	Urban	Rural	Pooled	Urban	Rural
Access	−0.0047 **	−0.1162 ***	−0.0042 **	−0.0047 **	−0.1184 ***	−0.0043 *
Log of Income	0.1026 ***	0.1510 ***	0.1107 ***	0.0978 ***	0.0658	0.0977 **
Occupation of Head == Agriculture	−0.5038 ***	−0.3241 ***	−0.5036 ***	−0.5071 ***	−0.3600 ***	−0.5125 ***
Occupation of Head == Labor	−0.5144 ***	−0.4556 ***	−0.5325 ***	−0.5173 ***	−0.4885 ***	−0.5403 ***

Table 2. *Cont.*

Variables	OLS-Estimates			IV-Estimates		
	Pooled	Urban	Rural	Pooled	Urban	Rural
Age	−0.0163 **	0.0088	−0.0238 ***	−0.0162 **	0.0150	−0.0236 ***
Age × Age	−0.00009	−0.0002	0.0002	0.00009	−0.0002	0.0002
Sex	0.1839 ***	0.1718 **	0.2155 ***	0.1829 ***	0.1639 *	0.2129 ***
Married	−0.0567 **	−0.0772	−0.0629 **	−0.0563 **	−0.0657	−0.0621 **
Excellent Health	0.1852 **	0.2507	0.1789 **	0.1872 **	0.2529	0.1852 **
Good Health	0.1436 *	0.2003	0.1422 *	0.1451 *	0.1923	0.1469 *
Privileged Caste	0.3192 ***	0.3468 ***	0.2916 ***	0.3211 ***	0.3951 ***	0.2963 ***
Indigenous Caste	−0.0425	−0.0186	−0.0558 *	−0.0416	0.0168	−0.0535 *
Central Region	−0.0904 ***	0.0342	−0.1366 ***	−0.0898 ***	0.0588	−0.1357 ***
Western Region	0.0041	0.0625	0.0042	0.0045	0.0832	0.0052
Mid-Western Region	−0.1357 ***	−0.0622	−0.1409 ***	−0.1361 ***	−0.0481	−0.1424 ***
Far Western Region	−0.0975 **	0.1178	−0.1072 ***	−0.0976 **	0.1325	−0.1076 ***
Hill Belt	−0.1215 ***	−0.1281	−0.1339 ***	−0.1219 ***	−0.1289	−0.1352 ***
Mountain Belt	−0.1533 ***	0.1151	−0.1463	−0.1543 ***	0.1085	−0.1489 ***
Time	−0.0654 ***	−0.0842	−0.0673 ***	−0.0609	−0.0029	−0.0555
Constant	1.8761 ***	0.7058	1.9884 ***	1.9202 ***	1.3934	2.1094 ***
Observations	4423	528	3709	4423	528	3709
R^2	0.3069	0.3060	0.3092	0.3069	0.2922	0.3089
First Stage coefficient[2] on the instrument (household labor hours)				0.5001 ***	0.3974 ***	0.5842 ***
First Stage F-Statistic				97.55	11.52	81.61
Durbin–Wu–Hausman Chi-sq Statistic				0.0171	0.3232	0.1053

Source: Calculated by authors; Standard Errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

5.2. Pooled Sample

The results for the pooled sample show that human capital rises with an increase in household income and better health status if the person is male and upper caste. Human capital falls with an increase in travel time to school, our access variable. The harder the access to schooling, the smaller the human capital is. What happens if the time to reach school is cut by half? From a plausible policy perspective, if the amount of time could be reduced from the overall sample mean of 125 min (2.0879 h) to 62.5 min, human capital could increase by 0.294 or 12.2% on average. This is a substantial improvement with high economic significance.

Human capital is negatively associated with farming (or "other labor") as the occupation of the household head, or if the person is from the far western or mid-western development region, or from mountains and hills.

Household income is an important economic variable. A 10% increase in household income is associated with a 0.0102 unit increase in human capital or by 0.425%. While the impact of an increase in household income on human capital formation is positive and significant, the magnitude of its impact is rather small. Among other variables, a typical male has 0.184 more units of human capital than a typical female, which indicates the existence of significant potential, going forward, for lowering male–female educational differences in Nepal. In terms of self-reported health status, an individual with excellent or good health will likely have a larger human capital by 0.1852 and 0.1436 units, respectively, than the rest of the population in the pooled sample.

A lower level of human capital is associated with an individual whose household head is a farmer or with an unskilled manual worker compared to one whose household head has an occupation in the service sector. An older male possesses less human capital than a younger one, and a male in an even older category has even less. A married individual has less human capital than a single one, and a person from the privileged class has more human capital than one from the indigenous and *Dalit* classes. All these results match our expectations.

In terms of the development regions, the average human capital is lower in three of the five development regions in the nation and higher in the western development region in Nepal than in the eastern, the reference region, though human capital was higher for all five regions in 2011 than in 2004. This is also brought out clearly in Table 2 by the coefficients of the time variable (dummy for 2011), which are positive for urban and rural samples and hence, for the nation as a whole.

5.3. Urban and Rural Subsamples

The rationale for sample stratification by sector (urban vs. rural) is to see whether there are significant differences in the way various factors affect human capital in urban and rural areas. The hypothesis is that the urban–rural split may impinge upon the slope as well as the shift of the regression line with respect to each variable. As Table 2 shows, we find the education access variable has an impact on human capital that is significantly higher in the urban areas than in rural (see Table 2).

In the urban sample, if the average value of access (travel time to school) is reduced by one-half (0.17 h or 10 min), human capital increases by 1.16 units or 42% of the urban average. This impact is large and is statistically significant at the 1% level.

Likewise, in the rural sector, if the travel time is reduced by half to 1.22 h (73 min), the rural human capital is predicted to increase by 0.343 units or 15% of the rural average (2.33). This impact is significant at the 5% level. A big increase in access to schooling is required for a larger impact on human capital in rural areas. This is because the average time to reach a school in the rural areas is 146 min compared to the urban time of 20 min. If Nepal were to achieve equality in average travel time to school between the two sectors at the current urban average of 20 min, the rural human capital would increase by 0.59 units or by 25.45%.

The different results pertaining to the access and several other explanatory variables in the case of rural and urban samples have important implications for Nepal. Children from higher-income households tend to have lower dropout rates compared to those from poorer households. This naturally widens inequality in human capital accumulation. Human capital is also greater for households who have service-oriented occupations compared to those in agriculture or those who have unskilled manual jobs. The impact of most of the other control variables is similar in urban and rural samples, except that in the urban sector, the results are insignificant for some control variables (see Table 2). The results show that central, mid-western, and far western regions have, as expected, lower levels of human capital compared to the eastern region, whereas the western region has attained the highest level of human capital per capita.

5.4. Decomposition

A discussion of human capital's relationship with socioeconomic factors leads us to take a closer look at how much those factors account for intersectoral or interregional differences in our human capital index (HCI). We present those results in Table 3. The top panel shows the predicted values of urban, rural, and national HCI at the corresponding means of the independent variables. The rural–urban difference (rural minus urban) in HCI is −0.3065 or 12.7% of the average HCI for the entire sample and is significant at the 1% level. The results of the Blinder–Oaxaca decomposition, also in Table 3, indicate relative proportions of the rural–urban gap in human capital that can be attributed to differences in the variable characteristics and differences in the variable parameters.

Table 3. Blinder–Oaxaca decomposition of differences in human capital between rural and urban using OLS regression.

Variables	Coeff.	Rob. SE	95% CI		Coeff.	Rob. SE	95% CI	
Group_1 (Rural)	2.2926	0.0111	2.2708	2.3144				
Group_2 (Urban)	2.5991	0.0293	2.5418	2.6565				
Difference (rural–urban)	−0.3065 ***	0.0319	−0.3691	−0.2439				
Explained		−0.1826						
Unexplained		−0.1238						
	Explained				**Unexplained**			
Access	−0.0061	0.0027	−0.0113	−0.0009	0.0267	0.0184	−0.0094	0.0628
Socioeconomic Characteristics	−0.1506	0.0168	−0.1835	−0.1177	−0.5494	0.5226	−1.5736	0.4748
Demographic Characteristics	−0.0504	0.0090	−0.0682	−0.0327	−0.7876	0.6173	−1.9974	0.4222
Geographic Characteristics	0.0245	0.0092	0.0065	0.0425	−0.0963	0.1353	−0.3614	0.1689
Constant					1.2827	0.8398	−0.3633	2.9287

Description of Variables. Access: Measured in terms of the average travel time to primary and secondary schools. Socioeconomic Characteristics: Income and Occupation. Demographic Characteristics: Age, Age x Age, Sex, Caste, Health Status, and Marital status. Geographic Characteristics: Development reasons such as Eastern, Western, Mid-Western, Far Western; and the ecological regions such as Hill, Mountain, and Terai. Source-Calculated by authors. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The negative and positive components, as explained and unexplained categories in Table 3, offset each other to a large extent. The net result is the unadjusted difference in human capital between rural and urban populations. As noted earlier, this difference equals −0.3065, which serves as the base for analyzing its decomposition. The explained component is the part of the outcome differential that is accounted for by rural–urban differences in the predictor variables or in their estimated parameters. The unexplained part can be attributed to the differences in unobserved variables (Jann 2008).

According to the results in Table 3, adjusting rural variables such as access to schooling to their levels in urban areas can eliminate the human capital gap by reducing the rural disadvantage by 0.1826, which will leave a negative gap of 0.1238 as unexplained. When adjusted by sectoral averages ("endowments"), the coefficients with a positive sign indicate a widening of the gap, while coefficients with a negative sign indicate a closing of the gap. Looking at the "explained" block of results in Table 3, most of the differences between rural and urban human capital relate to differences in socioeconomic and demographic characteristics, followed by a difference in geographic characteristics. The socioeconomic and demographic characteristics help to close the gap, whereas the geographic factor helps to widen the gap.

Of our particular interest is how much of the rural–urban difference in human capital is due to differences in access to schooling. The results indicate that access-to-school explains almost 0.0061, i.e., (0.0061/0.1826) = 3.34% of the human capital gap, and this factor actually helps to close the gap. Similar calculations show that differences in socioeconomic factors explain 0.1506 (or 82.47%) and the difference in demographic characteristics explain another 0.0504 (or 27.60%) of the human capital gap. These factors help to close the rural–urban gap, whereas the difference in geographic characteristics, which explains 0.0245 (13.41%), actually works to widen it. Broadly, our key result is that a reduction in the difference in access to schooling, socioeconomic characteristics, and demographic characteristics will significantly lower the human capital discrepancy between rural and urban populations.

Differences in the coefficients, as opposed to the variable averages, suggest that the socioeconomic, demographic, and geographic characteristics help offset the gap by 1.433. However, the access to schooling and the constant term lead to an increase in the rural–urban difference by 1.309, as the signs of the coefficients are positive. Overall, the differences in estimated parameters help to bring the rural and urban sectors together in human capital accumulation.

To sum up, our decomposition strategy indicates that despite its importance, access to schooling helps to close the gap in human capital by only 3.34%. In contrast, the socioeconomic and demographic variables account for the majority of the difference. Thus, in addition to educational and road infrastructures, improvements in social and institutional setup are also needed to nudge the rural economy toward parity with the urban economy. The efforts toward school expansion made around 1990 (Giri and Shrestha 2017) have contributed to lowering the inequality in human capital formation between villages and cities.

This paper contributes to the analysis of rural–urban dualism in a key component of development, that is, in the level of human capital. Access to education infrastructure can open the doors to other economic opportunities by making people more *capable* (Sen 1985). Sen argues that the ability to exercise options correlates with a more prosperous life in terms of three essential factors—knowledge, long and healthy life, and access to resources, all of which are needed for a decent standard of living to contribute to people's wellbeing. Our emphasis on access to education and road infrastructure supports the capability argument for increased wellbeing. A renewed emphasis on schooling access and road infrastructure is likely to boost Nepal's rural development further.

A reduction in travel time to school can be achieved in two ways: (i) build schools close to where people live, or (ii) build road infrastructure such that reaching a school is quicker. A universal agreement exists in the literature on the importance of education for people to improve their human capital. In a popular quote, Coleman (1988) argues "human capital is created by changes in persons that bring about skills and capabilities that make them able to act in new ways". Education infrastructure facilitates those changes and makes the overall development of an individual easier. In Nepal, the government's share of investment in education has been falling. A reversal of the policy on education infrastructure should go together with programs that are geared towards socioeconomic and demographic improvements.

5.5. Robustness Analysis

To check the robustness of our findings, we use instrumental variable regressions to re-estimate our parameters and to calculate the predicted rural–urban difference in human capital. After controlling for socioeconomic, demographic, and geographic factors, the impacts of access to schooling on human capital formation are still significant in IV-based regression models (both for the pooled sample and the urban and rural sub-samples). The direction of impacts in all three regression models is also the same. For instance, the partial impacts of access under OLS are -0.0047, -0.1162, and -0.0042, respectively, for the pooled, urban, and rural samples. Under our IV model, those effects are, respectively, -0.0047, -0.1184, and -0.0043. In both cases, all these impacts are statistically significant (see Table 2).[3]

We also find that the predicted differences in human capital formation are once again the same in the OLS and IV models and are highly significant, as indicated by Tables 3 and 4. The same is true of the explanatory power of the two models. As the size and the direction of the impacts in all three models are similar, and the predicted differences in human capital are equal and statistically significant, our findings from the OLS model appear relatively robust.

Table 4. Blinder–Oaxaca Decomposition of Differences in Human Capital between Rural and Urban using IV Regression.

Variables	Coeff.	Rob. SE	95% CI		Coeff.	Rob. SE	95% CI	
Group_1 (Rural)	2.2926	0.0111	2.2708	2.3144				
Group_2 (Urban)	2.5991	0.0293	2.5418	2.6565				
Difference (rural–urban)	−0.3065 ***	0.0319	−0.3691	−0.2439				
Explained			−0.1609					
Unexplained			−0.1456					
	Explained				**Unexplained**			
Access	−0.0072	0.0027	−0.0125	−0.0018	0.0285	0.0147	−0.0003	0.0574
Socioeconomic Characteristics	−0.1103	0.0154	−0.1405	−0.0800	1.1314	0.4435	0.2621	2.0007
Demographic Characteristics	−0.0589	0.0090	−0.0766	−0.0413	−1.1474	0.6414	−2.4045	0.1097
Geographic Characteristics	0.0155	0.0090	0.0022	0.0331	−0.3437	0.1409	−0.6198	−0.0676
Constant					0.1855	0.7298	−1.2448	1.6159

Description of Variables. Access: Measured in terms of the average travel time to primary and secondary schools. Socioeconomic Characteristics: Income and Occupation. Demographic Characteristics: Age, Age x Age, Sex, Caste, Health Status, and Marital status. Geographic Characteristics: Development reasons such as Eastern, Western, Mid-Western, Far Western; and the ecological regions such as Hill, Mountain, and Terai. Source-Calculated by authors. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

5.6. Limitations and Further Research

One of the important determinants of human capital is the individual level of skill. An excellent robustness check would be to perform an analysis including such a variable in the model. The lack of skill data is a limitation of this paper. However, with better individual level data on hand, a potential extension for further research would be to follow Gould and Moav (2016) and estimate the wage equation to generate residuals that can be used as a proxy for individual skill in a human capital model. Additionally, this paper highlights one of the three commonly used approaches to calculate HCI. However, along with the outcome, cost, and income approaches, there are authors such as Graham and Webb (1979) who also point out the need to consider the depreciation of human capital in those models. Future research could explore this idea further.

6. Conclusions

With 80% of the national population, Nepal's rural economy is afflicted with a large deficiency in human capital relative to its urban areas. We examine the supply side and demand side factors that account for the rural–urban difference in human capital based on large survey data. In particular, we focus on the role of education infrastructure as well as the socioeconomic characteristics of the sample households. Our empirical results suggest that access to schooling has a significant impact on the level of human capital, especially in the rural areas. This variable explains a modest portion of the rural–urban gap in human capital formation. A direct implication of this finding is that better access to schools in rural areas will help close this discrepancy. In terms of the size of the effects, the socioeconomic and demographic variables dominate. In particular, much of the difference between rural and urban human capital relates to household income and the occupation type of the household. These socioeconomic variables, together with demographic variables (age, sex, caste, current health status, and marital status), help to close the rural–urban gap in human capital. On the other hand, geographic characteristics examined along development regions and ecological zones, which were set up by a relatively autocratic political system over fifty years ago, actually magnify the rural disadvantage. Our results reinforce the claim that an improvement in schooling access and road infrastructure is necessary, particularly in the vast rural sector of Nepal, if human capital development is to provide a greater contribution to national welfare.

Notes

1 A detailed description of this methodology and estimation techniques in STATA are provided in Jann (2008).
2 To maintain brevity, results for the first stage (with log of income as the endogenous variable) are presented only for the instrument (household labor hours), followed by appropriate test results.
3 In running the instrumental variable regressions, we performed the Durbin–Wu–Hausman (DWH) test of endogeneity using the 'ivendog' command in Stata, following Baum et al. (2003). The test results are reported at the bottom of Table 2. Based on the test results alone, we cannot reject the null hypothesis that the explanatory variables are exogenous. However, theory suggests that human capital and household income are simultaneously determined and are endogenous. To address this issue, we used household labor hours as an instrument for household income. Then, we tested the validity of the instrument by examining the F-statistic from the first stage regression (using the 'estat firststage' command in Stata), which are well above the threshold value of 10, suggesting that household labor hours are a valid instrument for household income. Because the instrument is valid, and the fact that we cannot deny the issue of endogeneity (at least theoretically), we report the instrumental variable estimation results as well.

References

Akee, Randall, William Copeland, E. Jane Costello, and Emilia Simeonova. 2018. How Does Household Income Affect Child Personality Traits and Behaviors? *American Economic Review* 108: 775–827. [CrossRef] [PubMed]
Azam, Mehtabul. 2010. A Distributional Analysis of Social Group Inequality in Rural India. *Journal of International Development* 24: 415–32. [CrossRef]
Banerjee, Abhijit, and Esther Duflo. 2011. *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty*. New York: Public Affairs.
Bansak, Cynthia, and Brian Chezum. 2009. How do remittances affect human capital formation of school-age boys and girls? *American Economic Review* 99: 145–48. [CrossRef]
Barro, Robert J., and Jong-Wha Lee. 1993. International Comparisons of Educational Attainment. *Journal of Monetary Economics* 32: 363–94. [CrossRef]
Baum, Christopher F., Mark E. Schaffer, and Steven Stillman. 2003. Instrumental variables and GMM: Estimation and testing. *The Stata Journal* 3: 1–31. [CrossRef]
Blinder, Alan. 1973. Wage Discrimination: Reduced Form and Structural Estimates. *Journal of Human Resources* 8: 436–55. [CrossRef]
Charlery, Lindy C., Matin Qaim, and Carsten Smith-Hall. 2015. Impact of Infrastructure on Rural Income and Inequality in Nepal. *Journal of Development Effectiveness* 8: 266–86. [CrossRef]
Chattopadhyay, Saumen. 2012. *Education and Economics: Disciplinary Evolution and Policy Discourse*. Oxford: Oxford University Press.
Chattopadhyay, Saumen. 2014. Education for Development: An Exploration of the Linkages in the Indian context. *Trade and Development Review* 7: 51–71.
Calderón, Cesar, and Alberto Chong. 2004. Volume and Quality of Infrastructure and the Distribution of Income: An Empirical Investigation. *Review of Income and Wealth* 50: 87–106. [CrossRef]
Coleman, James S. 1988. Supplement: Organizations and Institutions: Sociological and Economic Approaches to the Analysis of Social Structure. *American Journal of Sociology* 94: S95–S120. [CrossRef]
Cunha, Flavio, James J. Heckman, Lance Lochner, and Dimitriy V. Masterov. 2006. Interpreting the evidence on life cycle skill formation. In *Handbook of the Economics of Education*. Amsterdam: North Holland, Vol. 1, pp. 697–812.
Currie, Janet. 2009. Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development. *Journal of Economic Literature* 47: 87–122. [CrossRef]
Currie, Janet, and Douglas Almond. 2011. Human Capital Development Before Age Five. In *Handbook of Labor Economics*. Amsterdam: Elsevier, Vol. 4, pp. 1315–486.
De, Prabal K., and Dilip Ratha. 2012. Impact of remittances on household income, asset and human capital: Evidence from Sri Lanka. *Migration and Development* 1: 163–79. [CrossRef]

Gilbert, John, and Nilanjan Banik. 2012. Socio-economic Impacts of Regional Transport Infrastructure in South Asia. In *Infrastructure for Asian Connectivity*. Cheltenham: Edward Elgar, pp. 139–63.

Giri, Animesh, and Vinish Shrestha. 2017. The Effect of School Construction on Educational Outcomes among Females: Evidence from Nepal. *Development Journal of the South* 2: 1–23.

Gould, Eric D., and Omer Moav. 2016. Does high inequality attract high skilled immigrants? *The Economic Journal* 126: 1055–91. [CrossRef]

Gradin, Carlos. 2015. Poverty and Ethnicity among Black South Africans. *European Journal of Development Research* 27: 921–42. [CrossRef]

Graham, John W., and Roy H. Webb. 1979. Stocks and depreciation of human capital: New evidence from a present-value perspective. *Review of Income and Wealth* 25: 209–24. [CrossRef]

Handa, Sudhanshu. 2002. Raising primary school enrolment in developing countries: The relative importance of supply and demand. *Journal of Development Economics* 69: 103–28. [CrossRef]

Jacoby, Hanan G. 2000. Access to Markets and The Benefits of Rural Roads. *The Economic Journal* 110: 713–37. [CrossRef]

Jann, Ben. 2008. The Blinder-Oaxaca Decomposition for Linear Regression Models. *The Stata Journal* 8: 453–79. [CrossRef]

Jorgenson, Dale W., and Barbara M. Fraumeni. 1989. Investment in Education. *Educational Researcher* 18: 35–44. [CrossRef]

Kendrick, John W. 1994. Total capital and economic growth. *Atlantic Economic Journal* 22: 1–18. [CrossRef]

Leipziger, Danny, Marianne Fay, Quentin Wodon, and Tito Yepes. 2003. *Achieving the Millennium Development Goals: The Role of Infrastructure*. World Bank Policy Research Working Paper 3163. Washington: The World Bank.

Mejía, Daniel, and Marc St-Pierre. 2008. Unequal opportunities and human capital formation. *Journal of Development Economics* 86: 395–413. [CrossRef]

Nehru, Vikram, Eric Swanson, and Ashutosh Dubey. 1993. *A New Database on Human Capital Stock: Sources, Methodology and Results (No. 1124)*. Washington: The World Bank.

Oaxaca, Ronald. 1973. Male-Female Wage Differentials in Urban Labor Markets. *International Economic Review* 14: 693–709. [CrossRef]

Psacharopoulos, George. 1994. Returns to investment in education: A global update. *World Development* 22: 1325–43. [CrossRef]

Sapkota, Jeet B. 2014. *Access to Infrastructure and Human Development: Cross-Country Evidence*. JICA-RI Working Paper No. 70. Tokyo: JICA Research Institute.

Sen, Amartya. 1985. *Commodities and Capabilities*. Amsterdam: North-Holland Publishing.

UNDP (United Nations Development Program). 2015. 2015 Human Development Report. Available online: http://hdr.undp.org/sites/default/files/2015_human_development_report.pdf (accessed on 1 April 2016).

Weil, David. 2016. Economic Growth, 3rd ed. New York: Routledge.

Woessmann, Ludger. 2003. Specifying Human Capital. *Journal of Economic Surveys* 17: 239–70. [CrossRef]

World Bank. 2017. World Development Indicators. Available online: http://data.worldbank.org/data-catalog/world-development-indicators (accessed on 1 April 2020).

World Economic Forum. 2015. The Global Competitiveness Report 2015–2016. Available online: http://www.weforum.org/issues/global-competitiveness (accessed on 26 April 2016).