Essays on the Empirical Study of Migration, Intrahousehold Trade-offs, and Infrastructure Investments

by

Slesh Anand Shrestha

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in The University of Michigan
2012

Doctoral Committee:

Associate Professor Dean Yang, Chair
Professor Charles C. Brown
Professor David A. Lam
Associate Professor Susan M. Dynarski
Assistant Professor Raj Arunachalam

To my mother and my grandmother for inspiring me all my life.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

ABSTRACT

Essays on the Empirical Study of Migration, Intrahousehold Trade-offs, and Infrastructure Investments

by

Slesh Anand Shrestha

Chair: Dean Yang

Education and infrastructure are two important development strategies in low-income countries. They foster economic growth by improving labor productivity and facilitating market integration. Evaluating them poses a considerable challenge because schooling choices and infrastructure constructions are influenced by unobserved characteristics that bias OLS estimates. This dissertation uses appropriate natural experiment and instrumental variable strategies to study the effect of skilled emigration prospects on human capital investment, identify inter-sibling spillovers in education, and evaluate the impact of proximity to road on farm profits.

Chapter 1 focuses on a natural experiment that involves the recruitment of Nepali men of Gurkha ethnicity into the British Army to identify whether improved prospects for skilled emigration may stimulate human capital investment at home. While the recruitment originated during the British colonial rule in South Asia, a change in the education requirement for Nepali recruits in 1993 resulted in an exogenous, differential increase in their skilled versus unskilled emigration prospects. I use individual-level information on ethnicity, gender, and age to estimate positive effects on school at-

tainment of Gurkha men. I extend this analysis in chapter 2 to study its impact on education of Gurkha women and estimate the net trade-offs between siblings' human capital investments. While I find that eligible men responded to the rule change by raising their schooling by over one year, a 30 % increase over their average, part of the improvements in their education came at the expense of their female counterparts living in the same household, whose education decreased by 0.11 years.

In chapter 3, I estimate the benefits of having easier access to road on farm profits. I overcome endogenous selection in road placement by constructing an instrument based on a geographic feature. Because mountains in Nepal stretch in a north-south direction, the cheaper cost of constructing a north-south road relative to an east-west road to connect the district headquarters led to greater access for villages in north-south hinterlands relative to those in east-west hinterlands. I find that farmlands appreciate in value by 0.25 percent when the travel time to road decreases by 1 percent.

CHAPTER I

Human Capital Investment Responses to Skilled Migration Prospects: Evidence from a Natural Experiment in Nepal

1.1 Introduction

"Brain drain" describes the emigration of skilled workers, mainly from developing to developed countries. It hinders countries' long-run economic growth by depleting their scarce human capital. In 2000 more than 12 million individuals from developing countries with tertiary education, 20 percent of their educated workforce, were living in OECD countries.

A new literature in skilled emigration emphasizes the effect of future emigration opportunities on human capital accumulation. The reasoning behind these "brain gain" models is that migration prospects raise the expected returns to education and encourage individuals at home to increase their human capital investment. If enough skilled individuals eventually decide not to emigrate, it can lead to a net increase in human capital stock at home.[1] While these new theoretical predictions challenge

[1]Individuals could change their emigration decision in the future for various reasons. The most popular explanation put forward by earlier brain gain papers is that quotas set by immigration authorities in destination countries are binding. Hence, even if lots of individuals actively seek to emigrate in the future, most of them are forced to live in their home country. Other reasons

the negative view of skilled emigration, empirical evidence is scarce because migration prospects are endogenous. This paper overcomes the problem by identifying a natural experiment that exogenously and differentially raised emigration prospects for skilled versus unskilled individuals at home.

The experiment involves a change in the British Gurkha Army recruitment regulations. The British Gurkha Army is the unit of the British Armed Forces that is composed of Nepali men of Monglo-Tibeto origins. According to *Banskota* (1994), since 1857 the Monglo-Tibeto tribes have created a tradition of enlisting in the British Gurkha Army, and thereby they are collectively referred to as the Gurkha ethnic group.[2] Furthermore, the economic and anthropological data from Nepal indicate that this British colonial tradition still plays an important financial and cultural role in the Gurkha communities of Nepal.[3] In 1993 the British government changed the educational requirement for the British Gurkha applicants, from requiring no education to requiring a minimum of 8^{th} grade education. This change was instigated by the modernization of the British Army following the developments in Eastern Europe in the early 1990s and the growing use of technology in warfare, as indicated by discussions in the House of Commons.[4] The reform focused on reducing service manpower and increasing technological capabilities, which were complemented by improving training and education of soldiers. Because this change was part of the broader restructuring of the British Army, the timing and the rationale for this change is exogenous to social, economic, or political characteristics of Nepal. Therefore, it

why individuals change their emigration decision, include unanticipated future changes in individuals' socio-economic characteristics as well as changes in their preferences which affect their future emigration choices.

[2] All the remaining ethnicities in Nepal, excluding the Gurkha ethnic group, are referred to as the "non-Gurkha" ethnic group for the remainder of this paper.

[3] For detailed documentation for the cultural significance and the financial contribution of the British Gurkha among Gurkha communities of Nepal, see *Hitchcock* (1966); *Caplan* (1995).

[4] For detailed documentation of the strategic defense reviews presented to the House of Commons regarding the restructuring of the British Army in early 1990s, together with the debate and discussions that followed, see the Eighth Defense Review. http://www.parliament.the-stationery-office.co.uk/pa/cm199798/cmselect/cmdfence/138/13802.htm

provides a simple natural experiment where Gurkha men experienced an exogenous increase in skilled emigration prospect relative to unskilled emigration to a foreign labor market i.e. the British Army.

My empirical approach improves identification compared to the strategies used in previous studies. The problem in estimating the impact of skilled emigration prospects on education arises for two reasons: first, unobserved characteristics, such as cultural norms and values, affect both migration and schooling decisions; and second, an increase in human capital raises migration incentives by reducing the domestic wage for skilled workers. *Beine et al.* (2006) address this problem by instrumenting current migration prospects with the past emigrant stock. But the same time-invariant characteristics that affect migration prospects and necessitate the use of an instrumental variable strategy are also likely to have affected migration in the past, thus undermining the validity of their historical instrument. While *McKenzie and Rapoport* (2006) argue that historical migration rates within Mexico are the outcome of early 20^{th} century railroad networks, it could still be problematic if past migration led to a better assimilation of foreign ideas, such as value of education, or if individuals with similar characteristics came to live together in regions with better access to railroads. Therefore, the ideal experiment would require two groups of individuals from developing countries that are identical except that skilled emigration prospects exogenously increase for one of them.[5] The change in education requirement of the British Gurkha Army created such an experiment.

Using Nepal Census data from 2001, I identify individuals' ethnicity to determine whether they were affected by this rule change. Since age-invariant ethnic characteris-

[5] To my knowledge, the only previous study to use a historical event for identification is *Chand and Clemens* (2008). They use the unexpected coup in Fiji as a source of variation to study the effect of emigration prospects on education. However, in contrast to my experiment, emigration prospects for the treatment group did not change exogenously but instead they argue that the decline in economic opportunity at home due to political instability created greater incentive for them to emigrate. In response, they invested in education that increased their likelihood of successfully emigrating to countries such as Australia.

tics could affect education, I use their age at the time of the rule change as the second criteria to determine their exposure. Given the recruits must be between the age of $17\frac{1}{2}$ and 21 years old, only men who were 21 or younger in 1993 could be affected by the rule change. I refer to these cohorts as the "eligible" cohort and older cohorts as the "ineligible" cohort in this paper. The difference-in-difference strategy controls for all age-varying characteristics and age-invariant ethnic characteristics that could be correlated with education. As an alternative strategy, I use the synthetic control method developed by *Abadie and Gardeazabal* (2003) by constructing a comparison for the Gurkha ethnic group based on ineligible cohorts' ethnic characteristics. The results suggest that the change in the educational requirement induced Gurkha men of cohorts aged 6 to 12 at the time of the rule change to raise their education by 1.11 years and cohorts aged 13 to 21 by 0.39 years. The estimates are consistent across the two strategies, highlighting the robustness of my empirical findings. More importantly, the rule change did not increase migration rates of eligible Gurkha men and increase in schooling also occurred for eligible Gurkha men who had not emigrated by 2001, which means that the rule change increased the net human capital stock of eligible Gurkha men.

My findings show that skilled emigration can lead to a net increase in human capital at home, by highlighting a positive impact of migration prospects on human capital investment. The early literature in skilled emigration ignore this relationship and focus solely on the loss of human capital from emigration of skilled workers, thereby concluding that skilled emigration is detrimental to home country's human capital accumulation.[6] By assuming that productivity increases with greater concentration of skilled workers, the endogenous growth models of *Miyagiwa* (1991) and *Haque and Kim* (1995) also predict negative impacts on economic growth. However, *Mountford* (1997) and *Stark et al.* (1997) show that endogenizing human capital accumulation

[6] *Bhagwati and Hamada* (1974) propose the brain drain tax to compensate for the loss incurred by developing countries due to skilled emigration.

on migration prospects, could raise net human capital stock of low income countries and improve their economic growth if some skilled individuals eventually decide not to emigrate. Since the recruitment was limited to 300 individuals annually, majority of eligible Gurkha men did not join the British Gurkha Army. Therefore, my findings are consistent with the theoretical predictions and overturn the pessimistic outcomes of the earlier models.

While *Foster and Rosenzweig* (1996) and *Kochar* (2004) show that schooling choices are affected by educational returns at home, my paper extends their analysis beyond political boundaries into foreign labor markets. I develop a theoretical model of human capital and emigration decisions, highlighting the three important factors that drive my results. First, there is a significant increase in income through emigration; second, there is an educational requirement to emigrate; and third, the likelihood of successfully emigrating is low. These three factors, which propelled eligible Gurkha men to invest in education and led to a net increase in their human capital following the rule change, are also the main drivers of brain gain in other developing countries. *Beine et al.* (2006) argue that large income differences between developing and developed countries increase the perceived benefits of migration, and induce emigration in developing countries. According to *Docquier et al.* (2007), popular destination countries such as Australia, Canada, the United Kingdom, and the EU employ skill-biased immigration policy, thereby making skilled emigration the most feasible way for individuals from developing countries to emigrate. They also find that the skilled emigration rate in developing countries in 2000 was only 7%, despite an increase of 70% over the previous decade. In line with my findings, *Chand and Clemens* (2008) show that skill-selective points systems used by immigration authorities of Australia and New Zealand induced Fijians to invest in education, resulting in a net increase in the human capital stock of Fiji. Therefore, the underlying mechanisms of my unique natural experiment are consistent with economic forces affecting individuals' education and migration choices in many developing countries.

The rest of the paper is structured as follows: Section 2 describes the natural experiment. Section 3 presents a theoretical framework that forms the basis of brain gain models. Section 4 explains the empirical strategy and the data used for causal estimation. Section 5 presents the empirical results. Section 6 presents robustness for the identification strategy and Section 7 concludes.

1.2 Background

Nepal is a landlocked country surrounded by India on three sides and China to its north. Its geographical position historically made it a melting ground for people and cultures from both north and south of its border (*Shrestha*, 2001b). The 1996 National Living Standard Survey categorizes the population of Nepal into 15 ethnic groups. Out of them, the Gurkha ethnic group is comprised of 5 Monglo-Tibeto tribal groups– the Rai, Limbu, Gurung, Magar, and Tamang, who settled in the eastern and central hills of Nepal during the initial wave of migration from the north.

1.2.1 British Brigade of Gurkha

The British Brigade of Gurkha is the unit in the British Army that is composed of Nepali soldiers. Following the Anglo-Gurkha war (1814-1816), the British East India company and the Government of Nepal signed the treaty of Sugauli on March 4, 1816. The treaty transferred one-third of the territories previously held by Nepal to the British and allowed them to set up three Gurkha regiments in the British Indian Army.[7] The early recruits of the British Gurkha Army included ethnic groups such as the Rajput, Thakuri, Chetri, and Brahman, who migrated from the south and were closely associated with the ethnicities of India. In 1857, Indian soldiers serving in the British Indian Army led a mutiny against British rule in South Asia.

[7] For details of the Sugauli Treaty see *Rathaur* (2001).

Although the rebellion was eventually contained in 1858, the British became wary of Indian nationals serving in their army. *Rathaur* (2001) and *Caplan* (1995) argue that, as a result, the British stopped recruiting Nepali individuals belonging to the ethnicities that originated from India into the British Gurkha Army. According to *Rathaur* (2001) after 1857, the new Nepali recruits were mainly drawn from the Rai, Limbu, Gurung, Magar, and Tamang ethnic groups who, unlike other ethnicities in Nepal, had migrated from north and had no cultural or historical ties with India.[8]

This ethnicity bias in the recruitment of British Gurkha Army continues to exist to the present day as the majority of the current British Gurkha soldiers are comprised of these 5 Monglo-Tibeto tribes. Although the British government no longer uses ethnicity as a criteria for selection, the de facto ethnicity bias could be due to the lack of information on recruitment available for non-Gurkhas, or because non-Gurkhas are marginalized in the recruitment process as the first stage of selection is conducted by ex-Gurkha servicemen who are themselves from Gurkha ethnic group. Consequently, this British colonial tradition has evolved into an important cultural identity and lucrative economic opportunity for the individuals of the Gurkha ethnic group. The present value of the lifetime income from serving in the British Gurkha Army for 25 years is estimated to be $ 1,334,091.81.[9] This includes a starting salary of $ 21,000 and a lifelong annual pension of about $ 15,000 after retirement. According to *Caplan* (1995), remittances from Gurkha soldiers and pensions for ex-Gurkha soldiers were the country's largest earner of foreign currency until the recent development of tourism and other sources of migration. Although every year there are only 300 successful applicants, the pay and pensions of the servicemen are the major source

[8]The preference for Gurkha ethnic men is evident from the letter by the British Commanding Recruitment Officer to Colonel Berkely, the British Resident at Kathmandu, in the early 1900s in which he writes, "I first consider his caste. If he is of the right caste, though his physique is weak, I take him" (*Banskota*, 1994).

[9]See Table 1.10 for estimation of the lifetime income for Gurkha soldiers. Other non-financial benefits include the opportunity to permanently settle in the UK along with immediate family members.

of capital in most Monglo-Tibeto communities in the hills of Nepal, mainly because their alternative employment is limited to farming.[10] The financial benefits of the British Gurkha Army in these communities is evident from quotes from the Gurkha households documented by *Caplan* (1995) such as, "One of my boys has gone to the Army, we have only that hope."

Although the appeal of joining the British Gurkha Army is driven by economic benefits, it also brings cultural prestige to Gurkha communities. *Caplan* (1995) points out that Gurkha ex-servicemen as well as their wives are known in their villages by their titles of the British Army and acquire considerable reputations to become the new elite in their communities. *Hitchcock* (1966) reports that many Gurkha villages are named after the title of their highest ranking British Gurkha officer, such as "the Captain's village." These narratives highlight the social, political, and economic stature wielded by the British Gurkha Army in Gurkha communities of Nepal.

1.2.2 Natural Experiment: A Change in Education Criteria

Education is an important aspect of the British Gurkha recruitment. Starting from 1993, recruits must have completed at least 8 years of education.[11] Prior to this, however, no formal education was required to join the British Army and the selection criteria was strictly limited to physical examinations. This change in the education requirement was instigated by the larger restructuring of the British Army in the early 1990s. Following the end of the cold war, a series of defense reviews termed "Option for Change" was conducted by the UK Ministry of Defense in order to evaluate the role of its army in the post-cold war era. It focused on reducing defense expenditure spurred by the economic benefits of the "peace dividend"[12] and, consequently, led to

[10] The Defense Committee of the House of Commons in 1989 suggested that the annual salary of British Gurkha soldier was 100 times the average income in the hills from where they come from.

[11] In 1997, the education requirement was further increased to a minimum of 10 years of education.

[12] Peace dividend is a political slogan popularized by US president George H.W. Bush and UK Prime Minister Margaret Thatcher after the end of the cold war. It describes the reduction in defense spending undertaken by many western nations, including as the US and the UK, and the

a reduction in service manpower of the British Army by 18 percent. Furthermore, this reduction was accompanied by the emphasis on a flexible and modernized force. This was achieved by incorporating new technologies in weapon systems, communications, reconnaissance missions, and intelligence gathering, and by improving education and training of soldiers. The Option for Change outlined the use of technology in future warfare and the importance of education for soldiers who use it, and concluded that "strong defense requires military capability of fighting in a high-technology warfare; the aim is smaller forces, better equipped, and properly trained" (Eighth Defense Report, 1997). In fact, the trend towards educated soldiers had already begun in the US Army, as its recruits with a high school diploma increased by 30 percent in the late 1980s.[13] Hence, the increase in educational requirement for the British Gurkha was induced by the modernization of the British Army in response to the increasing role of technology and the political changes in Eastern Europe and, therefore, it was exogenous to the socio-political events in Nepal.

1.3 Theoretical Model

Becker's model of human capital views education as an investment, where individuals compare their costs to future benefits. The future benefits from investing in education is an increase in lifetime income earned domestically when there is no opportunity to emigrate. Given positive emigration prospects, however, the future benefits should additionally include expected increase in income earned abroad. Furthermore, the latter could be larger if either the wage rate per human capital is higher abroad, or income abroad at all levels of education is greater and education is required to emigrate, or both. The following theoretical model, based on *Docquier*

subsequent redirection of those resources into social programs and a decrease in tax rates.

[13]According to the Tenth Quadrennial Review of US Military Compensation, the recruits who scored better than the median in the Armed Forces Qualification Test (AFQT) increased by 10% in early 1990s.

et al. (2007), highlights these positive effects of higher skilled emigration prospects on expected returns to schooling and, consequently, on educational investment and net human capital stock.

Consider a small developing economy, where labor is an important factor of production and is measured in efficiency units. All individuals at birth are endowed with a unit of efficiency. They live for two periods, youth and adulthood. There is an education program which if opted into during youth, increases the individual's efficiency in adulthood to $h > 1$. Furthermore, the heterogeneity among individuals is highlighted by the differences in their cost of the education program, denoted by c, which has a cumulative distribution $F(c)$ and density function $f(c)$ defined on R^+. Suppose the domestic economy is perfectly competitive so that workers are paid their marginal product, denoted by w. In youth, uneducated workers earn w and educated workers earn $w - c$. In adulthood, individuals can choose to work abroad, where the wage rate per efficiency unit is $\hat{w} > w$ and the income premium is $I > 0$. In adulthood, uneducated workers can either earn $\hat{w} + I$ if they migrate or w if they don't. Likewise, educated workers can earn $\hat{w}h + I$ if they migrate and wh if they don't migrate. Individuals incur a fixed cost in adulthood if they attempt to emigrate, denoted by M. Let the probability of migration, denoted by p, be the same for both educated and uneducated workers. Suppose $(\hat{w} + I - w) \geq \frac{M}{p}$, which implies that all individuals would choose to emigrate.[14]

If individuals are risk neutral so that they choose their education to maximize

[14]The condition for all educated workers to choose to migrate is $\hat{w}h + I - wh \geq \frac{M}{p}$ and for uneducated worker, it is $\hat{w} + I - w \geq \frac{M}{p}$. These conditions imply that an increase in income due to migration should be greater or equal to the ratio of the cost and probability of migration. According to the UK Defense Committee, the annual salary of a British Gurkha soldier in 1989 was 100 times the average income in Nepal. Furthermore, the financial cost of applying for the British Gurkha is minimal as there is no application fee and the recruiters visit most Gurkha villages every year during the first stage of selection process. Moreover, the empirical estimate is interpreted as the average treatment effect on all age eligible Gurkha men, regardless of their future intention of applying to the British Gurkha.

lifetime income, then the condition for investing in education becomes:

$$w - c \ + (1-p)\ wh \ + p\ (\hat{w}h + I) \ - M \quad > \quad w \ + (1-p)\ w \ + p\ (\hat{w} + I) \ - M$$

and individuals will invest in education if and only if:

$$c \ < c_p \ \equiv w(h-1) \ + p\ (\hat{w} - w)\ (h-1) \tag{1.1}$$

The critical threshold c_p is increasing in the probability of migration p, which implies that migration prospects raise the expected return to education and induce more individuals in developing countries to invest in education. Furthermore, this incentive effect is larger, greater the international wage difference $(\hat{w} - w)$. The share of adult domestic workers who opted for education in their youth is given by $H_p = F(C_p)$.

Now, suppose the migration probability changes differentially across educated and uneducated workers.[15] In line with the change in the British Gurkha education requirement, the migration probability is assumed to be still p for educated workers but $\underline{p}$ for unskilled workers, where $\underline{p}$ equals zero. The probability p is assumed to be exogenously determined by the service requirement of the British Army independent of schooling decisions of individuals at home.[16] Uneducated workers remain at home and therefore earn domestic wage w in both periods. In contrast, educated workers earn $w - c$ in their youth and once adult they can earn either $\hat{w}h + I - M$ if they

[15]Prior to the rule change in 1993, both educated and uneducated workers had equal probability of joining the British Gurkha Army. Furthermore, since the selection criteria were solely based on physical attributes, education did not increase one's chance of getting selected. Following the rule change in 1993, however, only individuals with required education level could apply; whereas, individuals with less than 8 years of education could no longer apply to the British Gurkha.

[16]The probability p can be a decreasing function of $c_s(p)$ in (2), defining an implicit solution for p. The response to the rule change is partly determined by individual's expectation of how others would response to the rule change, which in turn affect their perceived p. The perceived and actual probability after the rule change will either stay the same or increase.

migrate or $wh - M$ if they don't.[17] The new condition for investing in education is:

$$w - c + (1-p)\, wh + p\,(\hat{w}h + I) - M \quad > \quad w + w$$

and individuals will opt for education if and only if:

$$c < c_s \equiv w(h-1) + p\,(\hat{w} - w)\, h + p\, I - M \tag{1.2}$$

The new critical threshold c_s is increasing in skilled emigration probability p, the difference in wage $(\hat{w} - w)$, and the foreign income premium I. If $I \geq \frac{M}{p}$, it implies that $c_s > c_p$ because individuals who could emigrate without education previously, are now prompted to invest in education in order to earn income premium abroad. To sum up, emigration prospects raise expected returns to education because of higher wage rate abroad, and skilled emigration relative to unskilled emigration prospects further increase expected returns to education because only skilled workers can emigrate and earn the higher income premium.

After $pF(C_s)$ fraction of workers migrate abroad in their adulthood, the share of educated workers who are unable to emigrate relative to all domestic adult workers is:

$$H_s \quad = \quad \frac{(1-p)\, F(C_s)}{1 - pF(C_s)}$$

and $H_s > H_p$ if and only if:

$$p < \widetilde{p} \equiv \frac{F(C_s) - F(C_p)}{F(C_s)\,(1 - F(C_p))} \tag{1.3}$$

$\widetilde{p} < 1$ denotes the critical threshold probability of emigration for which the incentive effect of brain gain exceeds the negative effect of brain drain. Given an extremely

[17] Similar to the previous case, assume $\hat{w}h + I - wh \geq \frac{M}{p}$, so that all educated workers would choose to migrate.

low likelihood of getting selected into the British Gurkha Army, the probability of skilled emigration through the Gurkha recruitment is likely to be below $\widetilde{p}$, suggesting that the rule change would create a net gain in human capital.

1.4 Identification Strategy

The change in the educational requirement in 1993 compelled new recruits to complete at least 8 years of education in order to be eligible for the British Gurkha and, thereby, increased their skilled emigration prospect relative to unskilled emigration. The theoretical analysis above suggests that this rule change would increase expected returns to education and induce individuals to invest in human capital. Furthermore, because the rule change was exogenous to the socio-economic characteristics of Nepal, the empirical strategy of comparing the education outcome of individuals who were affected to those who were not affected by the rule change gives an unbiased estimate of its effect on domestic schooling.

The individuals' exposure is jointly determined by their sex, ethnicity, and age. First, the British Gurkha, in contrast to the other regiments of the British Army, is exclusively made up of men; therefore, women were not affected. Second, because recruits must be between $17\frac{1}{2}$ and 21 years old, men who were 22 or older in 1993 were not affected by the rule change.[18] Third, considering most British Gurkha soldiers since 1857 have been Gurkha ethnic men, non-Gurkha men were also not affected. Hence, the effect of rule change on age eligible Gurkha ethnic men is identified via difference-in-difference estimation, comparing male education between eligible and ineligible cohorts, within Gurkha and non-Gurkha ethnic groups. The difference in education between the two cohorts in the Gurkha ethnic group could be correlated with the age-varying unobserved variables. Therefore, subtracting from this the co-

[18] As mentioned earlier, cohorts 22 and older are referred to as the "ineligible" cohort and cohorts 21 and younger are referred to as the "eligible" cohort.

hort difference in education for non-Gurkha ethnic men would net out all age-varying characteristics as well as age-invariant ethnic characteristics that could directly affect education. The identification assumption is that in absence of this rule change in 1993, the evolution of education outcomes of men between the two cohorts would not have systematically differed across Gurkha and non-Gurkha ethnic groups. Furthermore, the difference-in-difference estimate of female education between the two ethnic groups and cohorts serves as a false experiment to test this identification assumption.

The above identification strategy can be expressed using the following regression framework:

$$\begin{aligned} Y_{ikml} \;&=\; c + \alpha_{1k} + \beta_{1m} + \eta_{1l} + \gamma\,(\,T_{ik} * G_{im}\,) + \\ &\sum_j (\,P_{ij} * R_m\,) * \delta_j + \sum_j (\,P_{ij} * K_l) * \lambda_j + \epsilon \qquad (1.4) \end{aligned}$$

where Y_{ikml} is the education outcome for individual i of age k and ethnicity m, born in district l; α_{1k} is an age dummy for each k; β_{1m} is an ethnicity dummy for each m; η_{1l} is a district of birth dummy for each l; G_{im} is a dummy indicating whether individual belongs to the Gurkha ethnic group; T_{ik} is a dummy indicating whether the individual belongs to eligible cohort; P_{ij} is a dummy indicating whether individual is age j for $j \in \{age\ cohorts\}$; R_m is a vector of ethnicity-specific variables; and K_l is a vector of district-specific variables.

The above reduced form specification nets out any positive or negative externalities that affect both Gurkha and non-Gurkha ethnic groups. On one hand, higher school enrollment among age eligible Gurkha men could decrease quality of education, which in turn could negatively impact schooling. On the other hand, higher school enrollment could create positive peer effects, encouraging schooling among Gurkha and non-Gurkha ethnic men who have no intention of joining the British Gurkha. Hence, the coefficient from the above specification should be interpreted as the net effect of the rule change on age eligible Gurkha ethnic men.[19] Furthermore, since the information regarding individual's decision to apply for the British Gurkha are not available, the coefficient is also the average treatment effect from the rule change on all age eligible Gurkha men, regardless of their future intention of applying for the British Gurkha.

The identification strategy can also be generalized to examine the impact of the rule change for each birth-year cohort in the following regression framework:

$$
\begin{aligned}
Y_{iklm} \quad = \quad & c + \alpha_{1k} + \beta_{1m} + \eta_{1l} + \sum_x (P_{ix} * G_{im}) * \gamma_x + \\
& \sum_j (P_{ij} * R_m) * \delta_j + \sum_j (P_{ij} * K_l) * \lambda_j + \epsilon \qquad (1.5)
\end{aligned}
$$

Each γ_x can be interpreted as the effect of the rule change on Gurkha men of age x. Since men who were 22 and older were not affected, the coefficients γ_x should be 0 for $x \geq 22$. Additionally, the coefficients γ_x should increase as x decreases for $x < 22$. Younger age eligible Gurkha men were more likely to be enrolled in school at the time of the rule change and had more time to complete 8^{th} grade education, putting them in a better position to respond to the rule change compared to older age eligible men.

The data required for the above identification strategy are obtained from 2001 Nepal Census for individuals aged 6 to 44 in 1993. It also includes information for individuals who were living abroad in 2001, abating the concern for potential bias due to selective attrition.[20] The census data are supplemented with Nepal Living Standard Survey (NLSS) from 1996, which is a household sample survey with greater

[19]Given Figure 1.1 suggests that Gurkha ethnic group is concentrated in specific regions of Nepal, the externalities generated by the response to the rule change is also more likely to be experienced by Gurkhas rather than non-Gurkhas. In this case, the estimated effect of the rule change should be interpreted as an overall effect on eligible Gurkha men, with a combination of the net effect of the rule change and net externality among Gurkhas generated by the rule change.

[20]If the entire household moved abroad between 1993 and 2001, then the information on those individuals are not available. However, the propensity for the entire household to move abroad is extremely low in Nepal.

detail. Table 1.1 presents summary statistics for the 1,389,705 individuals from the 2001 Census and the 3,373 households from the 1996 NLSS. The averages for some socio-economic characteristics are provided for the entire sample as well as separately for the Gurkha and non-Gurkha ethnic groups. The Gurkha ethnic group comprises of 16 percent of the samples in both surveys. Panel A shows that the average level of education for the Gurkha ethnic group is 3.28, which is slightly lower than the non-Gurkha average of 4.24. Similarly, 18.2% of Gurkha individuals were born in urban districts compared to 34.1% of non-Gurkhas. Panel B suggests that around the time of the rule change, non-Gurkha households had better access to school facilities than Gurkha households and 46.9% of non-Gurkha households were living in poverty compared to 48.5% of Gurkha households. Figure 1.1 shows the map of Nepal with the distribution of Gurkha ethnic group across 75 regional districts. It suggests that most Gurkha households live in the northern central region and north east corner of Nepal and predictably, the three British Gurkha recruitment centers are also located within these regions.

1.5 Results

The identification strategy can be illustrated with a simple difference-in-difference table between the eligible and ineligible cohorts in the Gurkha and non-Gurkha ethnic groups. Table 1.2 compares educational attainment of Gurkha and non-Gurkha men who were not affected by the rule change (age 22 - 28) to those who were affected, either cohort aged 6 to 12 or cohort aged 13 to 21. I use eligible cohort aged 6 to 12 as the preferred cohort of analysis because this younger eligible cohort is most likely to be enrolled in primary school in 1993 and also have enough time to change their education in line with the new rule by the time they apply to the British Gurkha. On the contrary, the ability of older eligible men aged 13 to 21 to respond to the rule change is determined by the years of education that they have had completed in 1993.

For example, a Gurkha men of age 20 would only be able to successfully respond to the rule change, if he had at least 7 years of education in 1993. Given the data on their education in 1993 are not available, the older eligible cohort includes Gurkha men some of whom were affected by the rule change and others who were not.

In both ethnic groups, average education increased over time; but it increased more in the Gurkha ethnic group. The simple difference-in-difference estimation shows that Gurkha men of younger eligible cohort (aged 6-12) completed an average of 1.2 more years of education. This is significantly difference from zero at the 1% level. Panel B shows that Gurkha men of older eligible cohort (aged 13-21) also raised their education by 0.28 years, which is less compared to younger eligible cohort but also expected due to the reasons discussed earlier. The two estimates are large in magnitude especially for younger eligible cohort with an increase in education of 32% over the ineligible cohort. The positive impact of the rule change indicates that the British Gurkha constitutes an attractive foreign labor market opportunity for Gurkha men. Furthermore, it highlights the role of skilled emigration prospects on increasing returns to education among individuals who might otherwise have limited opportunity to benefit from education in the domestic labor market. This is especially true for Gurkha recruitment as *Caplan* (1995) notes that most of the potential recruits come from rural villages of Nepal and if not for the British Gurkha Army their best alternative source of income is farming.

Tables 1.3 present the difference-in-difference analysis by estimating coefficient γ in equation (1.4). The specification in column 1 controls for age dummies and ethnicity dummies and the specification in column 2 additionally controls for district of birth dummies. Figure 1.1 shows that Gurkha ethnic groups are concentrated in the northern central region and north east corner of Nepal. If time-varying regional characteristics are correlated with education, it could bias the above estimates. I control for differential evolution of geographic regions in columns 3, 4, and 5 by including interactions of age dummies and district of birth dummies, for all ages and districts. The specification in columns 4 and 5 also include interactions of age dummies and district-level characteristics– total number of primary and secondary teachers in 1994. Moreover, the specification in column 5 controls for additional time-varying ethnic characteristic by including the interaction of age dummies and ethnicity-level variable measuring the travel time to school, obtained from 1996 NLSS. The errors in all specifications are clustered at the ethnicity level.

The estimates in Table 1.3, column 1 suggest that an increase in educational requirement for the British Gurkha raised the education among younger eligible cohort by 1.19 years and older eligible cohort by 0.42 years, and both estimates are statistically significant at the 1% level. More importantly, controlling for various time-varying regional and ethnic characteristics do not change the magnitude and the statistical significance of the estimates for both eligible cohorts, which makes it unlikely that the results are driven by time-varying characteristics that are correlated with education. The estimates in column 5, which includes all the controls mentioned earlier, suggest that Gurkha men from younger eligible cohort raised their education by 1.11 years and older eligible cohort by 0.39 years.

The above results rely on the assumption that in absence of the rule change, the difference of educational outcomes between the eligible and ineligible cohorts would not have systematically differed across Gurkha and non-Gurkha ethnic groups. Table 1.4 presents a series of control experiments that compare educational attainment between cohorts and ethnic groups that were not affected by the rule change and therefore, in contrast to the results in Table 1.3, should produce difference-in-difference estimates of zero. Panel A compares education of ineligible cohort aged 22 to 28 with another ineligible cohort aged 29 to 35 across Gurkha and non-Gurkha men. The estimate of coefficient λ in column 5 is 0.01 and not statistically different from zero at the conventional levels. The control experiment in panel B considers cohort aged 22 to 28 and cohort aged 38 to 44, so that the age difference between the two ineligible cohorts is consistent with the experiment in panel A of Table 1.2, in which the age difference between the younger eligible cohort and ineligible cohort is 9 years. The estimates are also not statistically different from zero in all specifications. While Gurkha women might not be an appropriate control group due to potential intrahousehold spillovers from an increase in education among Gurkha men, nevertheless, panel C compares education outcome for females aged 6 to 12 with 22 to 28, in Gurkha and non-Gurkha ethnic groups. Based on table 1.4, column 5 the estimate of the effect of the rule change on Gurkha women of younger eligible cohort is -0.09 and not statistically different from zero. Given that all the estimates of coefficient λ are not statistically different from zero, the results from table 1.4 support the validity of the identification assumption and suggest that the increase in education for age eligible Gurkha men in Table 1.3 is likely caused by the change in the educational requirement for the British Gurkha recruitment.

Table 1.5 shows the effect of the rule change for each eligible birth-year cohort by estimating γ_xs in equation (1.5) for $6 \leq x \leq 21$. The comparison group consists of ineligible cohort aged 22 to 28. In all specifications, the estimated effect is statistically significant at the 5% levels for Gurkha men 15 years or younger. The results in column 5 suggest that the rule change raised education for Gurkha men aged 15 by 0.69 years, aged 12 by 0.95 years, and aged 6 by 1.38 years. Furthermore, in line with the natural experiment, the effect of the rule change increases with younger age due to the reasons discussed in the earlier section. If the results are driven by the response to the rule change, the estimated effects would decrease with age for Gurkha men of eligible cohort and be zero for all ineligible birth-year cohorts. I test this hypothesis by estimating γ_xs in equation (1.5) for $6 \leq x \leq 35$. The control group comprises of men aged 36 and 37. The estimates of γ_xs are plotted in Figure 1.2. γ_xs fluctuate around zero and statistically insignificant for all $x \geq 22$ and increase as age decreases for

$x \leq 21$, providing further support for the internal validity of the natural experiment.

The rule change induced eligible Gurkha men who had no formal education at the time of the rule change, to enroll in school for the first time. Table 1.6 presents the effect at the extensive margin, by estimating the coefficient γ in equation (1.4) for younger eligible cohort, where the dependent variable is a dummy indicating years of education completed greater than zero. The results in column 5 suggest that the proportion of young eligible Gurkha men with at least 1 year of education increased by 10 percentage points. Given 51% of age ineligible Gurkha men have no formal education, the rule change induced 19.5 percent of young eligible Gurkha men who would not have received any formal education in the absence of the rule change, to enroll in school. In comparison, *Schultz* (2004b) estimates that the Mexican Progresa Program induced 10 percent of individuals who had no prior education to enroll in school by reducing educational cost by as much as 75%.[21]

Individuals who were induced to enroll in school by the rule change and who had already enrolled prior to the rule change, were further promoted to raise their education because the new rule required 8 years of education. I estimate the impact at different education levels by estimating the difference in differences in the cumulative distribution function of education between young eligible and ineligible cohorts across Gurkha and non-Gurkha men who have at least one year of education. Figure 1.3 depicts the estimates of γ^ss from equation (1.4), with a dummy dependent variable indicating the level of education completed equal to or greater than s, for each s = 2 to 15.[22] Among Gurkha men of young eligible cohort with at least one year of education, the share of those with 5 or more years (primary education) increased by 3 percentage points, 8 or more years (the requirement cutoff) increased by 6 percentage

[21]These two results might not be directly comparable as PROGRESA started from higher enrollment base and targeted poorest students. Because of these reasons, increasing schooling might have been harder to achieve in the case of PROGRESA.

[22]The error terms in these 14 seemingly unrelated regression equations (SURE) are correlated. In Figure 1.3, the 95 percent confidence intervals for each γ^ss are adjusted for cross-equation error correlation.

points, and 13 or more years (tertiary education) increased by 9 percentage points.

The large impact at upper end of the education distribution is particularly significant, given *Jensen* (2010) points out that in developing countries a combination of costs, low family income, and credit constraints provides a relatively greater hindrance to secondary schooling compared to primary education as it requires a longer term and more costly investment. For example, while 67% of Nepali boys in 1996 were enrolled in primary school, the net enrollment rate in lower secondary level (6-8 years) was merely 23% (1996 NLSS). Additionally, the positive impact on education above 8 years could be due to the further increase in the British Gurkha educational requirement from 8 to 10 years in 1997, or because higher education increased the likelihood of success with the introduction of English and mathematics exams in the selection process, or because eligible Gurkha men who completed 8 years of education to comply with the new rule continued into higher education. *Angrist and Imbens* (1995) find similar positive spillover effects in the United States, where the compulsory attendance laws induced a fraction of the sample to complete some college as a consequence of constraining them to complete high school.

In developing countries, socio-economic factors such as access to schools, costs, credit constraints, and family income, limit individuals from attending school even when they want to. I examine the effect of these factors on individuals' response to the rule change, by separately estimating equation (1.4) across different population characteristics. The results in columns 2 and 3 of Table 1.7 indicate that the effect of the rule change did not vary across districts with and without easy access to schooling. However, the difference in average travel time to school between the bottom and top quantile districts is only 0.3 hours, which reflects the emphasis put by the government on improving access to school in remote areas of Nepal. The results in columns 4 and 5 indicate that the impact of the rule change was smaller for individuals living in households that are involved in agricultural production. According to Nepal Living

Standard Survey from 2004, more than 10% of school-age children who were not enrolled in school indicated labor constraint in household work as the main cause of their absenteeism. Furthermore, more than 20% of the school absenteeism was caused by high financial cost of education. In oder to examine the role of poverty and credit constraints on schooling, columns 6 and 7 separately estimate the effect of the rule change across household income, using ownership of television set as a proxy for family wealth. In households that own a television set, Gurkha men of young eligible cohort raised their education by 1.28 years; whereas, their counterparts living in the household without a television set only raised their education by 0.76 years. The F-test suggests that estimates are statistically different at the 10% level. While these results should be interpreted with caution due to omitted variable bias, the stratified results, nevertheless, could be potentially informative given they document the role of poverty and credit constraints in limiting schooling in developing countries.

The majority of eligible Gurkha men would not join the British Gurkha Army because only 300 individuals are recruited every year. However, higher education could increase emigration rates of eligible Gurkha men through other channels besides the British Gurkha recruitment. I estimate coefficient γ in equation (1.4), with a dummy dependent variable indicating whether the individual was living abroad in 2001. Given everyone in the older eligible cohort would have had the chance to apply for the British Gurkha Army or pursue other emigration opportunities by 2001, I focus on this cohort to examine whether the rule change led to greater emigration among eligible Gurkha men. The estimates in Table 1.8, panel B suggest that there was no increase in migration rates among older eligible Gurkha men. The coefficients in all the specifications are zero and not statistically significant even at the 10% level. On the other hand, Table 1.9 estimates the effect of the rule change on education of eligible Gurkha men who had not emigrated by 2001. I estimate coefficient γ in equation (1.4) by only including individuals who were living in Nepal in 2001. The results in column 5 suggest that the rule change raised education of young eligible Gurkha men who had not emigrated by 1.14 years and older eligible Gurkha men of similar nature by 0.40 years. Both the estimates are statistically significant at the 1% level. Therefore, the results in Tables 1.8 and 1.9 together imply that the increase in educational requirement for the British Gurkha Army led to a net increase in the human capital stock of eligible Gurkha men.

1.6 Robustness

In the above empirical estimation, the non-Gurkha ethnic group may not be a valid comparison for the Gurkha ethnic group because ineligible Gurkha cohorts have significantly lower level of education than their non-Gurkha counterparts. To refute the possibility that the results could be driven by age-varying unobserved ethnic characteristics, I use a data-driven procedure developed by *Abadie and Gardeazabal* (2003) to construct a different comparison group. The new counterfactual– the synthetic Gurkha ethnic group– is the convex combination of all non-Gurkha ethnicities that most closely resemble the Gurkha ethnic group based on the education of age ineligible men. For each non-Gurkha ethnicity, the average years of education is calculated for each birth cohort and then ethnicity-weights are assigned to minimize the difference between education of Gurkha and synthetic Gurkha ethnic groups across ineligible cohorts aged 22 to 44.[23] Table 1.11 displays the weights of each non-Gurkha ethnicity in the synthetic Gurkha ethnic group.

Figure 1.4 depicts the years of education completed for Gurkha and synthetic Gurkha ethnic groups across birth cohorts aged 6 to 44. Education of the synthetic Gurkha ethnic group closely matches that of the Gurkha ethnic group for ineligible

[23] The ethnicity-weights are calculated from the minimization problem: Choose W to minimize $(X_G - X_N W)(X_G - X_N W)$, where $W = \{(w_1, ...w_J)'$ subject to $w_1 + ... + w_J = 1,\ w_J \geq 0\}$. X_G is a $(k\ X\ 1)$ vector of average years of education at each age ineligible birth cohorts for the Gurkha ethnic group, where $21 \leq k \leq 44$. X_N is a $(k\ X\ J)$ matrix with average years of education for k ineligible birth cohorts and J non-Gurkha ethnicities.

cohorts aged 22 to 44, suggesting that the eligible cohort of synthetic Gurkha ethnic group provide a close approximation to the eligible cohort of Gurkha ethnic group in the absence of the rule change. The difference in education between Gurkha and synthetic Gurkha ethnic groups for cohorts aged 6 to 21 could be interpreted as the effect of the increase in the British Gurkha education requirement. Figure 1.5 shows that education between Gurkha and synthetic Gurkha ethnic groups diverges considerably for eligible cohorts and the gap, depicted in Figure 1.5, becomes larger for younger cohorts, which is consistent with the results from the difference-in-difference estimation.

The results could have also been obtained entirely by chance. Following *Bertrand et al.* (2004), I iteratively apply the synthetic control method to all the non-Gurkha ethnicities to examine whether assigning treatment at random produces results of similar magnitude. In each case, the synthetic control is composed of the weighted combination of the remaining non-Gurkha ethnicities. Figure 1.6 displays the results of the placebo iterations for 10 non-Gurkha ethnicities. The faded lines show the gap in education between each non-Gurkha ethnicity and its corresponding synthetic version. The gap for Gurkha ethnic group, depicted by the dark line, is largest compared to any non-Gurkha ethnicities. More importantly, the education gap between Gurkha and synthetic Gurkha for eligible cohort is four times larger than the similar gap for ineligible cohort. Figure 1.7 shows that this is largest among all ethnicities. Given there are 11 different ethnicities, including Gurkha ethnic group, and thereby 11 different results, the probability of obtaining the largest effect for Gurkha ethnic group entirely by chance is $1/11 = 0.09$. Therefore, it is unlikely that the estimated effect of the rule change on age eligible Gurkha men could have occurred entirely by chance.

1.7 Conclusion

The change in the educational requirement for the British Gurkha Army in 1993 led to an exogenous and differential increase in skilled versus unskilled emigration prospects for Gurkha men of eligible cohort living in Nepal. Using a set of difference-in-difference and synthetic control strategies, I find that they responded to the rule change by raising their human capital investment. I also find that the rule change increased education for eligible Gurkha men who had not joined the British Gurkha or emigrated elsewhere by 2001. These two findings validate the theoretical predictions of the brain gain models: first, individuals' human capital investments are endogenous to their migration prospects; and second, when enough skilled individuals eventually decide not to emigrate, it leads to a net increase in human capital stock at home.

The underlying mechanism of my unique natural experiment is not different from economic factors influencing individuals' emigration and human capital decisions in many developing countries. Despite an extremely low chance of getting selected into the British Gurkha Army, Gurkha men were induced to invest in education following the rule change because of significant increase in income if they succeeded in joining the British Army. *Docquier et al.* (2007) point out that these two factors, widening international wage gaps and introduction of skilled-biased immigration policies, are the main reasons for a rapid growth of skilled emigration and for inducing human capital investment among potential emigrants in developing countries. Nevertheless, they find that in 2000 the skilled emigration rate in developing countries was only 7%.

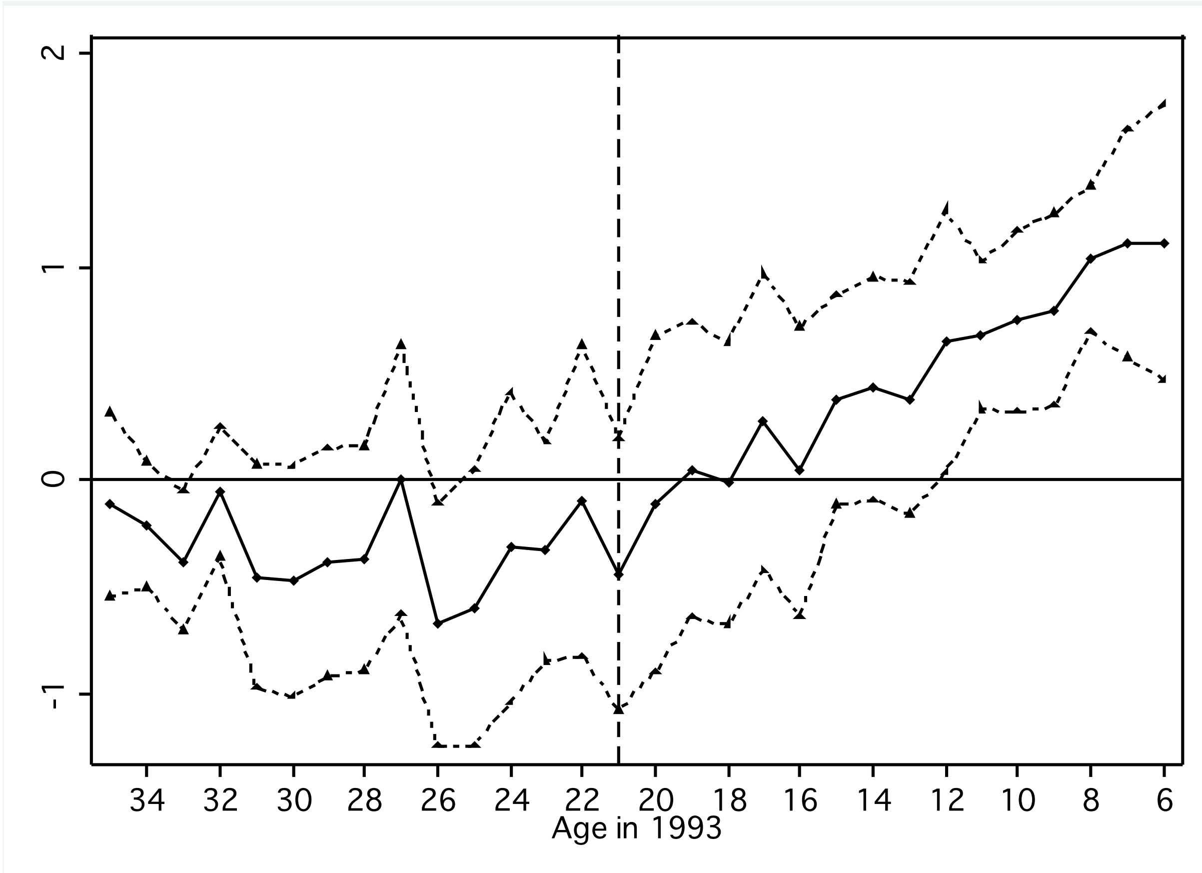
While the knowledge of the British Gurkha rule change was widespread, similar information about other foreign labor markets may not be as readily available. *Jensen* (2010) show that the lack of information regarding returns to education could lead to underinvestment in human capital. Therefore, it might require an efficient flow of information possibly through an active government intervention for individuals at home to know their educational returns abroad, and consequently, to increase their human capital investment.

An important implication of my findings is that low-income countries do not have to wait for improvements in their local productivity to stimulate human capital investment because high wages in developed countries can motivate individuals in developing countries to invest in education. While there is little doubt that low levels of schooling deter economic growth, *Schultz* (1975) show that returns to schooling are low in a stagnant economy, hinting at the possibility of a poverty trap. *Oyelere* (2009) argue that poor institutions and political instability, characteristics that are common across many developing countries, lead to low returns to education. Developing countries spend large sums of money on increasing human capital investment in order to overcome low returns and push themselves out of the poverty trap, if it exists. For example, Mexico's Progresa Program which provides cash incentive to increase school attendance, costs almost 0.2 percent of its GDP. Since skilled emigration prospects raise educational returns in developing countries, it could either replace expensive policy interventions like Progresa or complement these programs, making them more attractive to their potential recipients.

Interesting future research includes investigating the welfare impact on eligible Gurkha men who who could not join the British Gurkha Army. Increase in their education could raise their domestic earnings, improve their children's health outcomes, and promote long-term economic growth in their regions. Similarly, the rule change also created numerous positive and negative externalities on other populations that were not directly affected by it. First, an increase in education by Gurkha men could directly affect their peers' education decisions. On one hand, it decreases the quality of education by crowding out classrooms; whereas, on the other hand, greater class participation could lead to a positive learning experience for other classmates. This provides a useful experiment to investigate peer effects, which is an integral part of education research. Second, raising Gurkha men's education could also affect the education of their siblings, mainly female siblings who were not affected by the rule change. Given key socio-economic decisions including children's education are taken at household-level, investigation into the specific mechanisms governing this intra-household tradeoffs is important. It would allow for better evaluation of existing household interventions and development of more effective policies in the future.

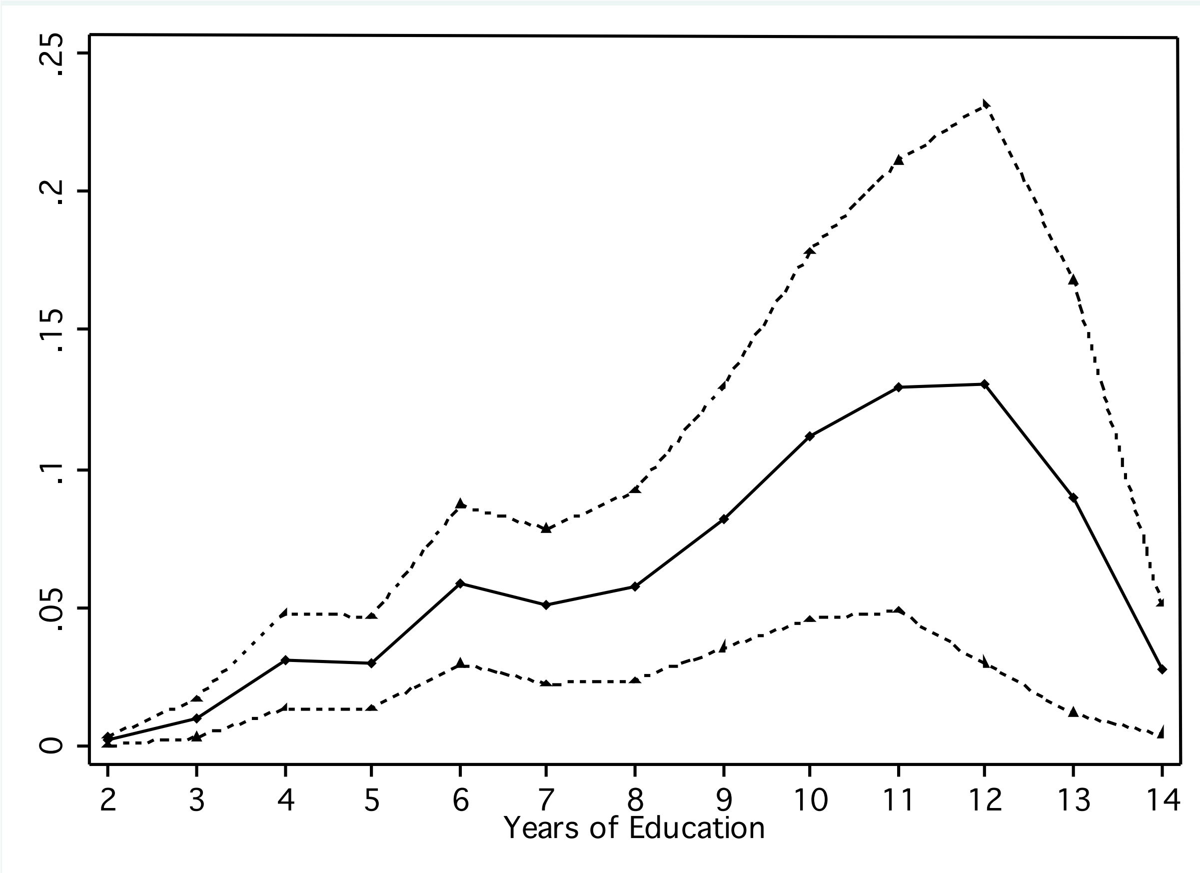
Figure 1.1: Map of Nepal with Concentration of Gurkha Ethnic Group and the British Gurkha Recruitment Centers

Figure 1.2: Effect of the Rule Change on Human Capital Investment of Gurkha Men at Each Birth Cohort (Identification Test)



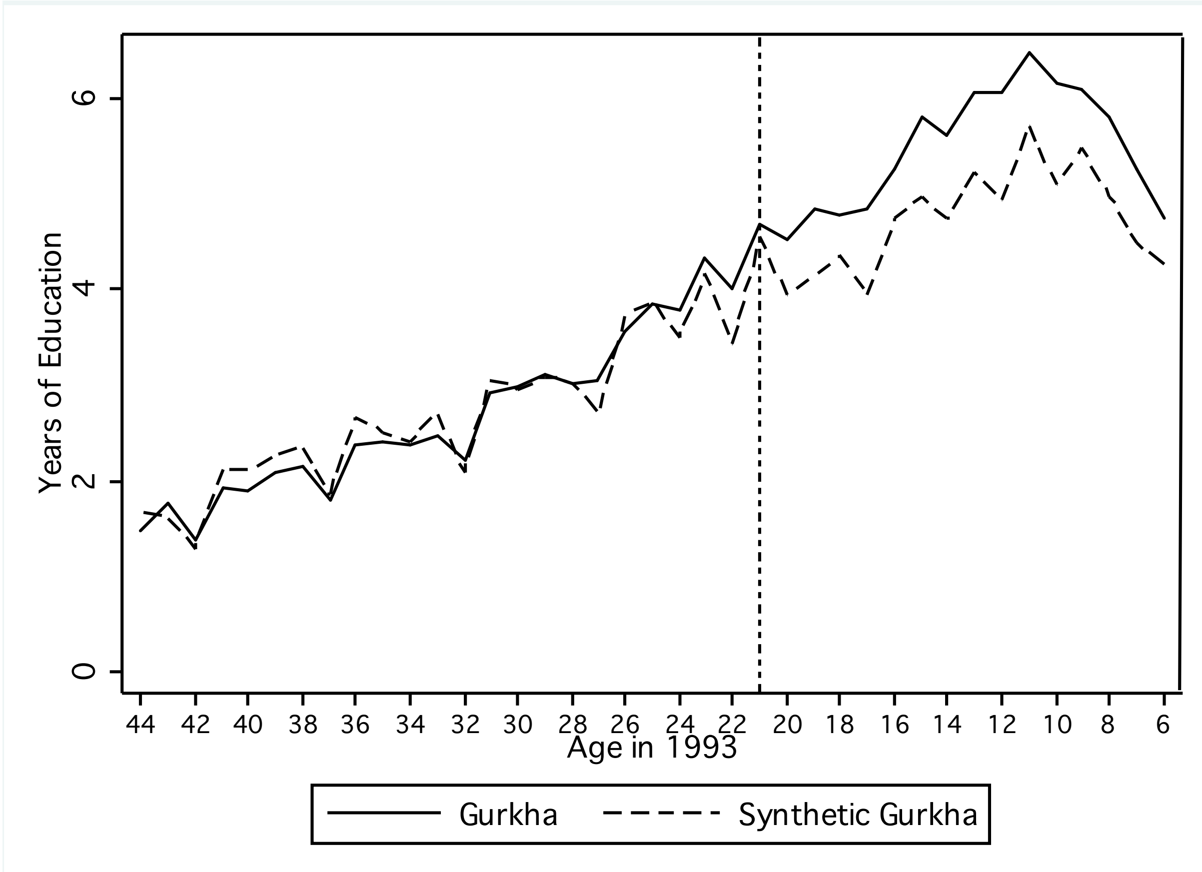
Notes: The figure above plots γ_xs for $6 \leq x \leq 35$ from equation (1.5). Since each γ_x estimates the effect of the rule change on Gurkha men of age x in 1993, γ_x should be zero for $x \geq 21$ and increase as x decreases for $x < 21$.

Figure 1.3: Difference in Differences in CDF (Estimated from Linear Probability Model) with 95-Percent Confidence Interval



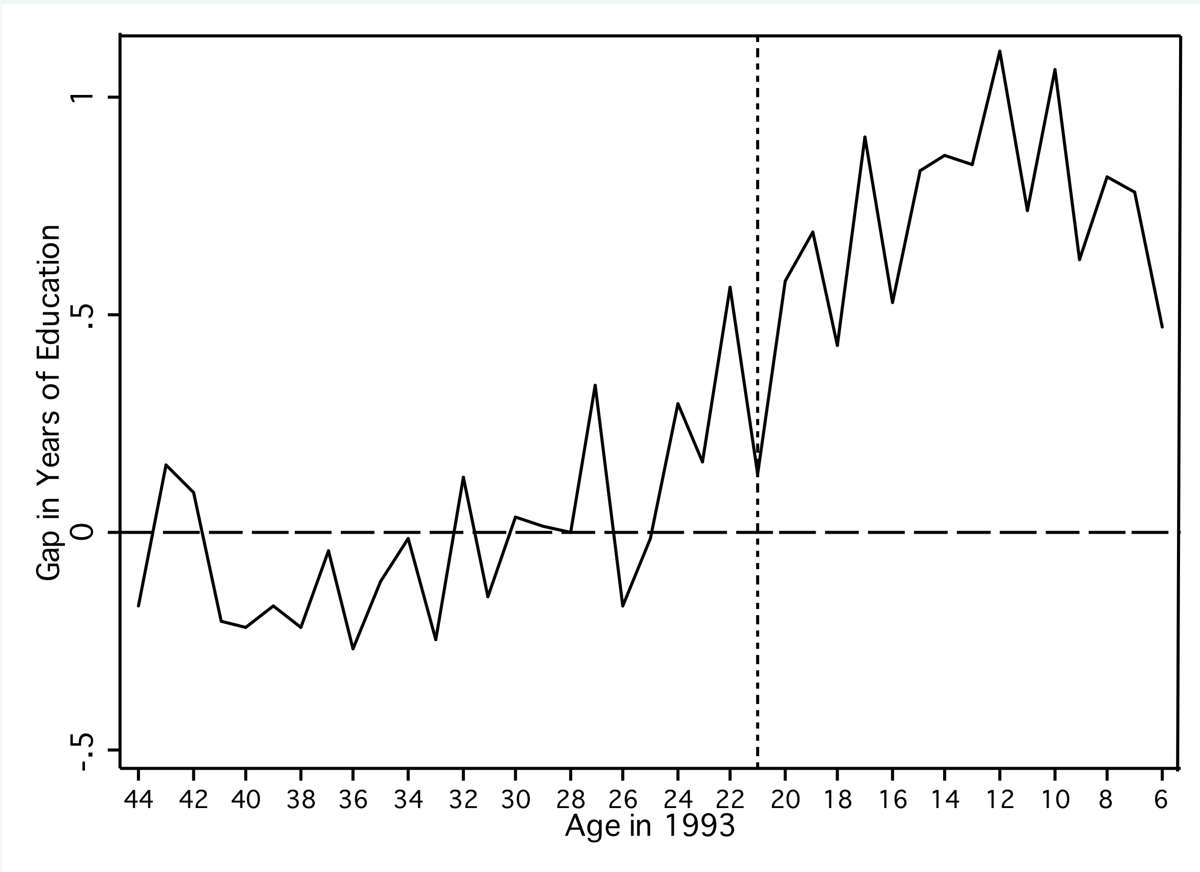
Notes: The figure plots γ^s estimated from equation (1.4) with a dummy dependent variable indicating the years of education completed greater than or equal to s, for each $s = 2$ to 14. The error terms in these 14 seemingly unrelated regression equations (SURE) are correlated. The 95 percent confidence intervals for each γ^ss are adjusted for cross-equation error correlation. The sample includes men from younger eligible cohorts aged 6 to 12 or ineligible cohort aged 22 to 28, with at least 1 year of education completed. Each γ^s indicate the impact of the rule change at the education level s among Gurkha men of younger eligible cohort with at least 1 year of schooling.

Figure 1.4: Comparison between Gurkha and Synthetic Gurkha



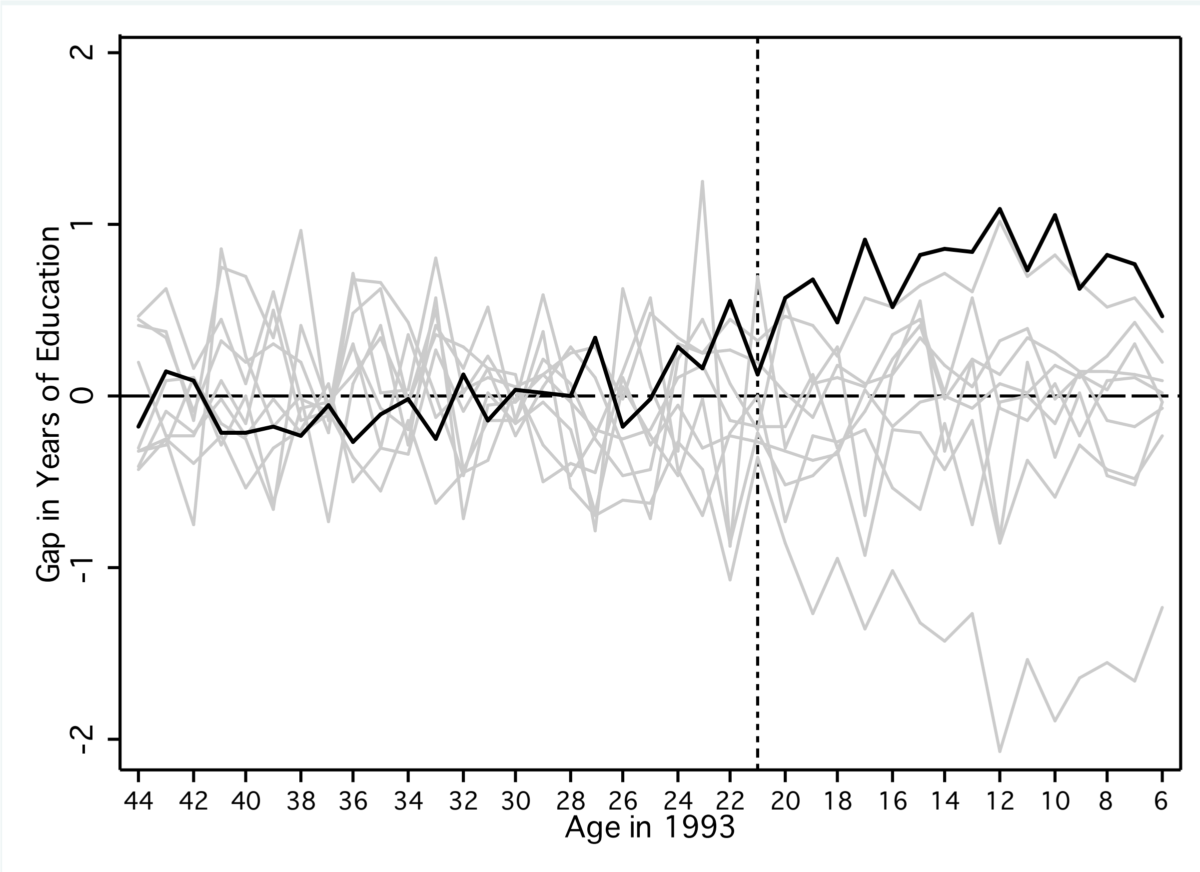
Notes: The graph above plots the average years of education completed as of 2001 at each birth cohort for Gurkha and synthetic Gurkha ethnic groups. The synthetic Gurkha is a weighted sum of all the non-Gurkha ethnicities. The weights are calculated to minimize the squared difference in average education of Gurkha and synthetic Gurkha ethnic groups across birth cohorts aged 22 to 44. Based on the mathematical algorithm provided by *Abadie and Gardeazabal* (2003), I choose W to minimize $(X_G - X_N W)(X_G - X_N W)$, where $W = \{(w_1, ...w_J)'$ subject to $w_1 + ... + w_J = 1,\ w_J \geq 0\}$. X_G is a $(k\ X\ 1)$ vector of average years of education at each age ineligible birth cohorts for the Gurkha ethnic group, where $21 \leq k \leq 44$. X_N is a $(k\ X\ J)$ matrix with average years of education for k ineligible birth cohorts and J non-Gurkha ethnicities.

Figure 1.5: Comparison between Gurkha and Synthetic Gurkha (Gap)



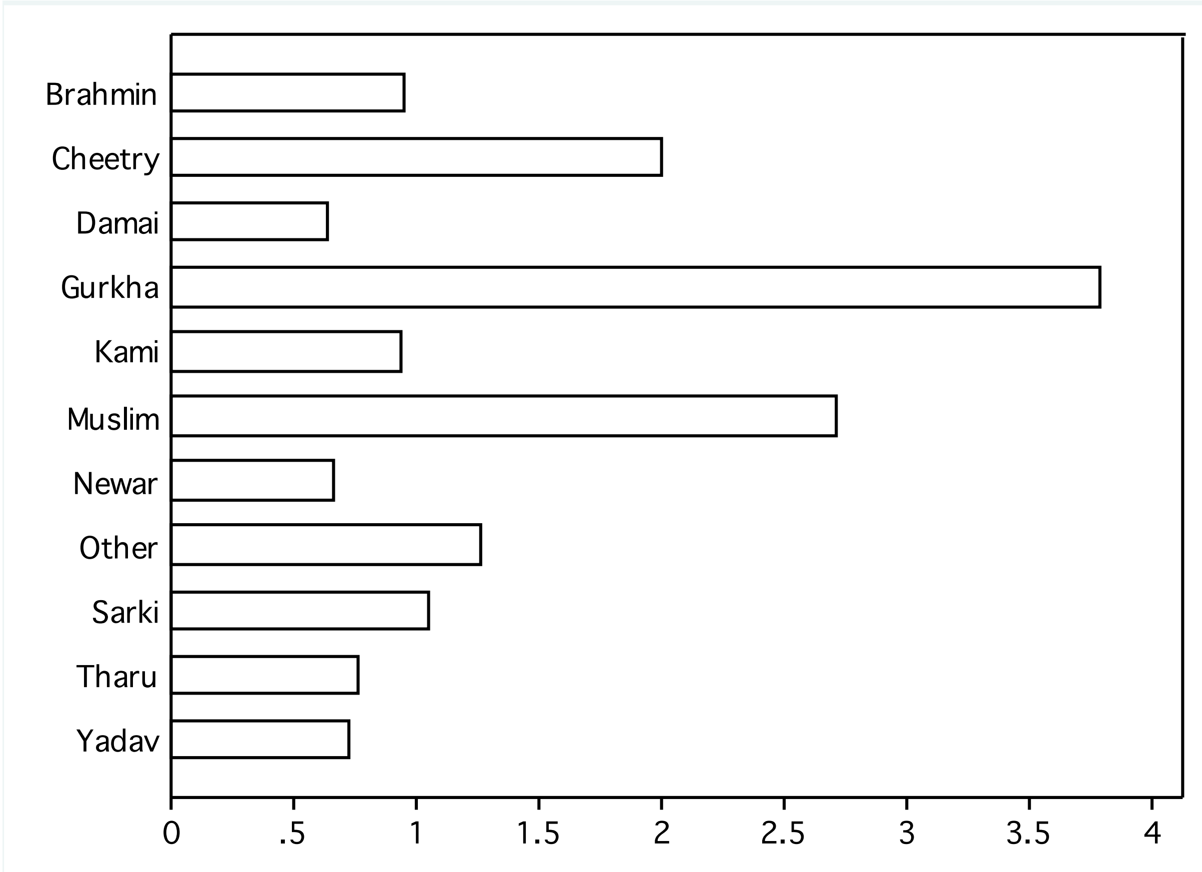
Notes: The figure plots the difference between average education of Gurkha and synthetic Gurkha ethnic groups for each birth cohorts aged 6 to 44, i.e. the difference between the education trends of the two groups from Figure 1.4.

Figure 1.6: Significance Test: Gap for Gurkha and 10 Placebo Gaps for Non-Gurkha Ethnicities



*Notes:*The figure plots the gaps same as Figure 1.5 for Gurkha ethnicity in the dark line and similar gaps for 10 non-Gurkha ethnicities in faded lines. For each non-Gurkha ethnicity, its synthetic counterpart is calculated by assigning weights to the remaining non-Gurkha ethnicities in order to minimize the squared difference in average education between the two groups across birth cohorts aged 22 to 44.

Figure 1.7: Significance Test: Ratio of Eligible and Ineligible Cohort Education Gap for Gurkha and Non-Gurkha Ethnicities



Notes: The figure shows the ratio of average difference in education between ethnicity and its synthetic counterpart for eligible and ineligible cohorts i.e. $\frac{(Avg\ Education\ Gap)_{Eligible\ cohort}}{(Avg\ Education\ Gap)_{Ineligible\ cohort}}$. This is largest for Gurkha ethnicity, which means that the probability of getting this result by chance is $1/11 = 0.09$.

Table 1.1: Descriptive Statistics

	Whole Sample	Gurkha	Non-Gurkha
Panel A: Individual Level Means			
2001 Census of Nepal			
Total Sample	1389705	245148	1144557
% of sample	-	17.6%	82.4%
Literacy Rate	55.2%	53.2%	55.6%
male	69.5%	66.9%	70.0%
female	41.3%	41.0%	41.4%
Level of Education	4.07	3.28	4.24
aged 6-12 in 1993	5.41	5.08	5.48
male	6.17	5.74	6.26
female	4.64	4.46	4.68
aged 13-21 in 1993	4.89	3.96	5.08
male	6.38	5.20	6.61
female	3.53	2.95	3.66
aged 22-28 in 1993	3.41	2.33	3.62
male	5.08	3.64	5.34
female	1.82	1.22	1.95
aged 29-37 in 1993	2.49	1.46	2.71
male	3.97	2.44	4.28
female	1.02	0.55	1.12
aged 38-44 in 1993	1.88	1.03	2.10
male	3.10	1.76	3.42
female	0.60	0.30	0.67
Percent of Population Born in Urban	31.3%	18.2%	34.1%
aged 6-12	34.7%	21.1%	37.8%
aged 13-21	32.7%	19.6%	35.4%
aged 22-28	30.3%	17.3%	32.9%
aged 29-37	27.8%	15.0%	30.5%
aged 38-44	25.6%	13.2%	29.0%
Panel B: Household Level Means			
1996 NLSS			
Total Sample	3373	544	2829
% of Sample	-	16.1%	83.9%
Household Size	5.59	5.27	5.65
Access to School	0.38 Hrs	0.54 Hrs	0.35 Hrs
Access to Paved Road	9.30 Hrs	14.45 Hrs	8.30 Hrs
Percent of Household in Poverty	33.5%	48.5%	46.9%

Table 1.2: Mean Education by Cohort and Ethnicity

	Level of Education Completed		
	Gurkha	Non-Gurkha	Difference
Panel A: Experiment 1			
Male aged 6 to 12 in 1993	5.74	6.26	-0.53
	(0.018)	(0.009)	(0.022)
Male aged 22 to 28 in 1993	3.64	5.34	-1.71
	(0.031)	(0.016)	(0.039)
Difference	2.10	0.92	1.18
	(0.034)	(0.017)	(0.042)
Panel B: Experiment 2			
Male aged 13 to 21 in 1993	5.20	6.61	-1.41
	(0.026)	(0.013)	(0.031)
Male aged 22 to 28 in 1993	3.64	5.34	-1.71
	(0.031)	(0.016)	(0.039)
Difference	1.56	1.27	0.29
	(0.041)	(0.020)	(0.049)

Notes: This table reports the mean education completed as of 2001 for men of different cohorts and ethnic groups. While Gurkha ethnic men of age 21 and younger were affected by the rule change, those who were 6 -12 years old were more likely to have been enrolled in school at the time of the rule change and, thereby, be in a better position to change their education. On the other hand, the ability of older eligible men aged 13 to 21 to respond to the rule change is determined by the years of education that they have had completed in 1993. Hence, this older eligible cohort includes Gurkha men some of whom were affected by the rule change and others who were not.
The standard errors are reported in parentheses.

Table 1.3: Effect of the Rule Change on Human Capital Investment of Eligible Gurkha Men

	(1)	(2)	(3)	(4)	(5)
Independent Var: Gurkha*Eligible					
Panel A: Experiment 1					
Eligible Cohort: Males aged 6 to 12	1.27**	1.19**	1.17**	1.20**	1.11**
Ineligible Cohort: Males aged 22 to 28	(0.340)	(0.342)	(0.364)	(0.386)	(0.342)
Obs: 325,876					
Panel B: Experiment 2					
Eligible Cohort: Males aged 13 to 21	0.45**	0.42**	0.42**	0.46**	0.39**
Ineligible Cohort: Males aged 22 to 28	(0.108)	(0.115)	(0.100)	(0.104)	(0.114)
Obs: 300,327					
Control Variables:					
Age Fixed Effects	Yes	Yes	Yes	Yes	Yes
Ethnicity Fixed Effects	Yes	Yes	Yes	Yes	Yes
District of Birth Fixed Effects	No	Yes	Yes	Yes	Yes
Age Dummies*District of Birth Dummies	No	No	Yes	Yes	Yes
Age Dummies*Rural Birth Dummy	No	No	Yes	Yes	Yes
Age Dummies*Primary School Teacher[a]	No	No	No	Yes	Yes
Age Dummies*Secondary School Teacher[b]	No	No	No	Yes	Yes
Age Dummies*Access to School[c]	No	No	No	No	Yes

Notes: This table reports the estimates of γ, the coefficient of interaction between Gurkha dummy and eligible cohort dummy from equation (1.4). The dependent variable is the years of education completed as of 2001. Specifications correct for various time-varying geographic and ethnic characteristics that could be correlated with education and bias the estimate. Eligible Gurkha men aged 6 -12 years old were more likely to have been enrolled in primary school at the time of the rule change and, thereby, would have been in a better position to change their education compared to eligible men aged 13 to 21. Hence, I use Panel A as the preferred experiment for analysis.

[a]**Primary School Teacher** is district-level total number of primary school teachers in pubic schools in 1994.

[b]**Secondary School Teacher** is district-level total number of secondary school teachers in pubic schools in 1994.

[c]**Access to School** is ethnicity-level average distance-time to school in 1996 calculated using 1996 NLSS.

The standard errors are reported in parentheses and are adjusted for within-ethnicity correlation between individuals.

* indicates significance at the 5 percent level

** indicates significance at the 1 percent level

Table 1.4: Effect of the Rule Change on Human Capital Investment of Ineligible Gurkha Men and Eligible Gurkha Women (Falsification Tests)

	(1)	(2)	(3)	(4)	(5)
Independent Var: Gurkha*Eligible					
Panel A: Control Experiment 1					
Eligible Cohort: Males aged 22 to 28	0.23	0.23	-0.02	0.03	0.01
Ineligible Cohort: Males aged 29 to 35	(0.213)	(0.224)	(0.186)	(0.176)	(0.166)
Obs: 214,315					
Panel B: Control Experiment 2					
Eligible Cohort: Males aged 22 to 28	-0.09	-0.06	-0.48	-0.43	-0.62
Ineligible Cohort: Males aged 38 to 44	(0.241)	(0.428)	(0.261)	(0.467)	(0.466)
Obs: 192,046					
Panel C: Control Experiment 3					
Eligible Cohort: Females aged 6 to 12	0.57	0.54	0.12	0.17	-0.09
Ineligible Cohort: Females aged 22 to 28	(0.408)	(0.407)	(0.345)	(0.332)	(0.351)
Obs: 333,055					
Control Variables:					
Age Fixed Effects	Yes	Yes	Yes	Yes	Yes
Ethnicity Fixed Effects	Yes	Yes	Yes	Yes	Yes
District of Birth Fixed Effects	No	Yes	Yes	Yes	Yes
Age Dummies*District of Birth Dummies	No	No	Yes	Yes	Yes
Age Dummies*Rural Birth Dummy	No	No	Yes	Yes	Yes
Age Dummies*Primary School Teacher	No	No	No	Yes	Yes
Age Dummies*Secondary School Teacher	No	No	No	Yes	Yes
Age Dummies*Access to School	No	No	No	No	Yes

Notes: This table reports the estimates of γ, the coefficient of interaction between Gurkha dummy and eligible cohort dummy from equation (1.4). All three experiments estimate the effect on individuals who were not affected by the rule change. Panel A and B estimates the effect on Gurkha men of cohorts, aged 22 to 28. In panel B, the age difference between the two cohorts aged 22 to 28 and aged 38 to 44 is the same as in panel A of Table 1.4. Panel C estimates the effect on Gurkha women of eligible cohort aged 6 to 12. None of the Gurkha women were affected because the recruitment of British Gurkha is limited to men. The lack of significant results in these three control experiments provide support for the validity of the identification assumption used in Table 1.4 that the difference in education between the cohorts would have been same across Gurkha and non-Gurkha ethnic groups in the absence of the rule change.

* indicates significance at the 5 percent level

** indicates significance at the 1 percent level

Table 1.5: Effect of the Rule Change on Human Capital Investment of Eligible Gurkha Men at Each Birth Cohort

	(1)	(2)	(3)	(4)	(5)
Gurkha*Age 21	-0.30	-0.32	-0.18	-0.14	-0.20
	(0.162)	(0.167)	(0.119)	(0.120)	(0.149)
Gurkha*Age 20	0.21	0.19	0.16	0.18	0.18
	(0.115)	(0.132)	(0.119)	(0.122)	(0.189)
Gurkha*Age 19	0.31*	0.30*	0.32**	0.36**	0.34**
	(0.104)	(0.110)	(0.070)	(0.062)	(0.079)
Gurkha*Age 18	0.20	0.17	0.25*	0.30*	0.27
	(0.118)	(0.121)	(0.094)	(0.093)	(0.142)
Gurkha*Age 17	0.86**	0.85**	0.54*	0.59*	0.40
	(0.192)	(0.188)	(0.184)	(0.181)	(0.188)
Gurkha*Age 16	0.23	0.22	0.31	0.34	0.36
	(0.186)	(0.182)	(0.155)	(0.163)	(0.190)
Gurkha*Age 15	0.50*	0.49*	0.64**	0.69**	0.60**
	(0.196)	(0.203)	(0.171)	(0.166)	(0.181)
Gurkha*Age 14	0.74**	0.71**	0.70**	0.72**	0.56**
	(0.138)	(0.136)	(0.161)	(0.180)	(0.182)
Gurkha*Age 13	0.65**	0.57*	0.65**	0.69**	0.68**
	(0.200)	(0.209)	(0.205)	(0.218)	(0.194)
Gurkha*Age 12	1.13**	1.09**	0.91**	0.95**	0.80**
	(0.223)	(0.220)	(0.214)	(0.227)	(0.213)
Gurkha*Age 11	0.94**	0.86*	0.94**	0.98**	0.81**
	(0.293)	(0.298)	(0.281)	(0.303)	(0.265)
Gurkha*Age 10	1.23**	1.12**	1.01*	1.06**	0.98**
	(0.224)	(0.229)	(0.276)	(0.297)	(0.256)
Gurkha*Age 9	1.08**	1.01*	1.06*	1.09*	1.04*
	(0.377)	(0.373)	(0.371)	(0.394)	(0.351)
Gurkha*Age 8	1.39**	1.27**	1.31**	1.33**	1.25**
	(0.409)	(0.413)	(0.405)	(0.428)	(0.418)
Gurkha*Age 7	1.51**	1.41**	1.38*	1.39*	1.34*
	(0.454)	(0.466)	(0.491)	(0.513)	(0.462)
Gurkha*Age 6	1.39*	1.30*	1.38*	1.38*	1.36*
	(0.601)	(0.598)	(0.557)	(0.579)	(0.541)
Control Variables:					
District of Birth Fixed Effects	No	Yes	Yes	Yes	Yes
Age Dummies*District of Birth Dummies	No	No	Yes	Yes	Yes
Age Dummies*Rural Birth Dummy	No	No	Yes	Yes	Yes
Age Dummies*Primary School Teacher	No	No	No	Yes	Yes
Age Dummies*Secondary School Teacher	No	No	No	Yes	Yes
Age Dummies*Access to School	No	No	No	No	Yes

Notes: All specifications include age and ethnicity fixed effects. The standard errors are adjusted for within-ethnicity correlation between individuals.

Table 1.6: Effect of the Rule Change on School Enrollment of Young Eligible Gurkha Men (Using Linear Probability Model)

	Obs	(1)	(2)	(3)	(4)	(5)
Extensive Margin						
Dep Var: Dummy Indicating Years of Education >0						
	325,876	0.12**	0.12**	0.10**	0.10*	0.10*
		(0.033)	(0.034)	(0.034)	(0.035)	(0.034)
Control Variables:						
Age Fixed Effects		Yes	Yes	Yes	Yes	Yes
Ethnicity Fixed Effects		Yes	Yes	Yes	Yes	Yes
District of Birth Fixed Effects		No	Yes	Yes	Yes	Yes
Age Dummies*District of Birth Dummies		No	No	Yes	Yes	Yes
Age Dummies*Rural Birth Dummy		No	No	Yes	Yes	Yes
Age Dummies*Primary School Teacher		No	No	No	Yes	Yes
Age Dummies*Secondary School Teacher		No	No	No	Yes	Yes
Age Dummies*Access to School		No	No	No	No	Yes

Notes: This table reports the estimates of γ, the coefficient of interaction between Gurkha dummy and eligible cohort dummy, from equation (1.4) with a dummy dependent variable indicating the years of education completed greater zero. The sample includes men from younger eligible cohorts aged 6 to 12 or ineligible cohort aged 22 to 28. Because prior to the introduction of educational requirement in 1993 no formal education required to join the British Gurkha Army, the above coefficients estimate the effect of the rule change at the extensive margin. The standard errors are reported in parentheses and are adjusted for within-ethnicity correlation between individuals.

* indicates significance at the 5 percent level

** indicates significance at the 1 percent level

Table 1.7: Effect of the Rule Change on Human Capital Investment of Young Eligible Gurkha Men (Differential Effect)

Whole	Distance to School[a]		Household Enterprise[b]		Household Income[c]	
Sample	≤Median	>Median	Ag Land	No Ag Land	Owns TV	No TV
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1.17**	1.16*	1.18**	0.86*	1.34**	1.29**	0.76**
(0.364)	(0.416)	(0.347)	(0.284)	(0.391)	(0.320)	(0.269)
325,879	181,159	144,717	130,698	195,178	86,851	239,025

Notes: This table reports the estimates of γ, the coefficient of interaction between Gurkha dummy and eligible dummy from equation (4). The dependent variable is the years of education completed as of 2001. The sample includes men of eligible cohort aged 6 to 12 and ineligible cohort aged 22 to 28. All the columns use specification from column 3 in Table 1.4, which includes ethnicity dummies, age dummies, district of birth dummies, age dummies*district of birth dummies, and age dummies*rural birth dummies.

[a]**Distance to School:** Sample divided into two categories based on average travel time to school in 1996: individuals born in (1) districts with average travel time to school ≤ median district-level travel time and (2) districts with average travel time to school ≥ median district-level travel time. The median district-level travel time to school is 0.36 hours.

[b]**Household Enterprise:** Sample divided into two categories based on ownership of agricultural land and livestocks in 2001.

[c]**Household Income:** Sample divided into two categories based on ownership of a television set in 2001.

The standard errors are reported in parentheses and are adjusted for within-ethnicity correlation between individuals.

* indicates significance at the 5 percent level

** indicates significance at the 1 percent level

Table 1.8: Effect of the Rule Change on Migration of Eligible Gurkha Men (Using Linear Probability Model)

	(1)	(2)	(3)	(4)	(5)
Dependent Var: Dummy Indicating Foreign Residence in 2001					
Panel A: Experiment 1					
Eligible Cohort: Males aged 6 to 12	-0.01	-0.01	-0.01**	-0.01**	-0.01**
Ineligible Cohort: Males aged 22 to 28	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)
Obs: 325,876					
Panel B: Experiment 2					
Eligible Cohort: Males aged 13 to 21	-0.00	-0.00	-0.00	-0.00	-0.00
Ineligible Cohort: Males aged 22 to 28	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
Obs: 300,327					
Control Variables:					
Age Fixed Effects	Yes	Yes	Yes	Yes	Yes
Ethnicity Fixed Effects	Yes	Yes	Yes	Yes	Yes
District of Birth Fixed Effects	No	Yes	Yes	Yes	Yes
Age Dummies*District of Birth Dummies	No	No	Yes	Yes	Yes
Age Dummies*Rural Birth Dummy	No	No	Yes	Yes	Yes
Age Dummies*Primary School Teacher[a]	No	No	No	Yes	Yes
Age Dummies*Secondary School Teacher[b]	No	No	No	Yes	Yes
Age Dummies*Access to School[c]	No	No	No	No	Yes

Notes: This table reports the estimates of γ, the coefficient of interaction between Gurkha dummy and eligible cohort dummy from equation (4). The dependent variable is the dummy variable indicating whether the individual is living abroad in 2001. Eligible Gurkha men, especially cohort aged 13 to 21, would have had the opportunity to apply for the British Gurkha Army by 2001 and also to seek other emigration opportunities. The coefficients are statistically insignificant for cohort aged 13 to 21, suggesting that the increase in education in response to the rule change did not increase their propensity to emigrate. If anything, the emigration is reduced for eligible Gurkha men aged 6 to 12, as suggested by the negative and statistically significant coefficients in Panel A.

[a]**Primary School Teacher** is district-level total number of primary school teachers in pubic schools in 1994.

[b]**Secondary School Teacher** is district-level total number of secondary school teachers in pubic schools in 1994.

[c]**Access to School** is ethnicity-level average distance-time to school in 1996 calculated using 1996 NLSS.

The standard errors are reported in parentheses and are adjusted for within-ethnicity correlation between individuals.

* indicates significance at the 5 percent level

** indicates significance at the 1 percent level

Table 1.9: Effect of the Rule Change on Human Capital Investment of Eligible Gurkha Men (Non-Migrants)

	(1)	(2)	(3)	(4)	(5)
Independent Var: Gurkha*Eligible					
Panel A: Experiment 1					
Eligible Cohort: Males aged 6 to 12	1.30**	1.22**	1.19**	1.22**	1.14**
Ineligible Cohort: Males aged 22 to 28	(0.341)	(0.344)	(0.366)	(0.388)	(0.344)
Obs: 320,592					
Panel B: Experiment 2					
Eligible Cohort: Males aged 13 to 21	0.45**	0.42**	0.42**	0.46**	0.40**
Ineligible Cohort: Males aged 22 to 28	(0.108)	(0.116)	(0.101)	(0.105)	(0.114)
Obs: 293,946					
Control Variables:					
Age Fixed Effects	Yes	Yes	Yes	Yes	Yes
Ethnicity Fixed Effects	Yes	Yes	Yes	Yes	Yes
District of Birth Fixed Effects	No	Yes	Yes	Yes	Yes
Age Dummies*District of Birth Dummies	No	No	Yes	Yes	Yes
Age Dummies*Rural Birth Dummy	No	No	Yes	Yes	Yes
Age Dummies*Primary School Teacher[a]	No	No	No	Yes	Yes
Age Dummies*Secondary School Teacher[b]	No	No	No	Yes	Yes
Age Dummies*Access to School[c]	No	No	No	No	Yes

Notes: This table reports the estimates of γ, the coefficient of interaction between Gurkha dummy and eligible cohort dummy from equation (4). The dependent variable is the years of education completed as of 2001. The sample includes only men who had not emigrated by 2001. Specifications correct for various time-varying geographic and ethnic characteristics that could be correlated with education and bias the estimate. Eligible Gurkha men, especially cohort aged 13 to 21, would have had the opportunity to apply for the British Gurkha Army by 2001 and also to seek other emigration opportunities. The positive impact on the education of eligible Gurkha men who had not emigrated suggests that the rule change led to a net increase in the human capital stock of eligible Gurkha men.

[a]**Primary School Teacher** is district-level total number of primary school teachers in pubic schools in 1994.

[b]**Secondary School Teacher** is district-level total number of secondary school teachers in pubic schools in 1994.

[c]**Access to School** is ethnicity-level average distance-time to school in 1996 calculated using 1996 NLSS.

The standard errors are reported in parentheses and are adjusted for within-ethnicity correlation between individuals.

* indicates significance at the 5 percent level

** indicates significance at the 1 percent level

Table 1.10: Lifetime Earnings of British Gurkha Soldier

Timeline	Income	Present Value of Income[a]
1. Recruitment into the British Gurkha at age 20[b]		
New Entrant	$20,880.00	$20,880.00
Rifleman (Level 1)	$26,037.00	$25,776.63
Rifleman (Level 2)	$28,068.00	$27,509.45
Rifleman (Level 3)	$31,000.00	$30,079.27
Rifleman (Level 4)	$33,320.00	$32,007.06
Rifleman (Level 5)	$36,842.00	$35,036.38
Lance Corporal (Level 1)	$38,634.00	$36,373.14
Lance Corporal (Level 2)	$40,407.00	$37,661.96
Lance Corporal (Level 3)	$42,224.00	$38,961.97
Lance Corporal (Level 4)	$44,256.00	$40,428.62
Lance Corporal (Level 5)	$44,256.00	$40,024.33
Corporal (Level 1)	$44,256.00	$39,624.09
Corporal (Level 2)	$44,286.00	$39,254.44
Corporal (Level 3)	$46,454.00	$40,764.36
Corporal (Level 4)	$47,539.00	$41,299.31
Corporal (Level 5)	$48,686.00	$41,872.80
Corporal (Level 6)	$49,694.00	$42,312.34
Corporal (Level 7)	$50,779.00	$42,803.81
Sergeant (Level 1)	$50,779.00	$42,375.77
Sergeant (Level 2)	$51,424.00	$42,484.90
Sergeant (Level 3)	$52,727.00	$43,125.78
Sergeant (Level 4)	$53,392.00	$43,232.99
Sergeant (Level 5)	$54,432.00	$43,634.36
Sergeant (Level 6)	$55,471.00	$44,022.58
Sergeant (Level 7)	$56,512.00	$44,400.24
Subtotal	$1,092,355.00	$955,946.58

Lifetime Earnings of British Gurkha Soldier (Continued)

2. Retirement from British Gurkha at age 45[c]		
(Between age 45 to 65, the pension is calculated under Early Departure Payments Structure)		
(i) Between 45 to 55:		
Highest Pensionable Salary*Years of Service*1/70*(50%+8.3335%)=		
56512*25*1/70*58.3335%=$11,773.37		
x 10=	$117,733.67	$87,562.87
1^{st} Lump Sum = 3 * Pension	$35,320.11	$27,472.74
Subtotal	$153,053.78	$115,035.61
(ii) Between 55 to 65:		
Highest Pensionable Salary*Years of Service*1/70*(75%)=		
56512*25*1/70*75%=$15,137.14		
x10=	$151,371.43	$101,815.75
Subtotal	$151,371.43	$101,815.75
(After age 65, the pension is calculated under AFPS 05)		
(iii) Between age 65 to 75[d]:		
Highest Pensionable Salary*Years of Service*1/70=		
56512*25*1/70=$20,182.85		
x10=	$201,828.57	$122,773.76
2^{nd} Lump Sum = 3 * Pension	$60,548.57	$38,520.11
Subtotal	$262,377.14	$161,293.87
Total	**$1,659,157.35**	**$1,334,091.81**

Notes: This table reports the lifetime income of a typical Gurkha soldier, including annual salary, benefits, and pensions.

[a] Present value of income is calculated using a discount factor of 0.99. If income in the third year of service is $28,068, then the present value of third year's income is 0.99^2*28,068 = $27,509.45

[b] Salary structure for British Gurkha soldier is based on "Rates of Pay as of April 2009" published by the UK Ministry of Defense

[c] Pension scheme is based on "Armed Forces Pension Scheme 05: Your Pension Scheme Explained" published by Service Personnel Policy (Pensions), The UK Ministry of Defense on January 2007.

Table 1.11: Ethnicity Weights in the Synthetic Gurkha

Ethnicity[a]	Weight
Cheetry	0.063
Brahmin	0.034
Tharu	0.094
Newar	0.045
Kami	0.157
Yadav	0.074
Muslim	0.095
Damai	0.130
Sarki	0.233
Other	0.076

Notes: This table reports the weights on each non-Gurkha ethnicities in the synthetic Gurkha ethnic group. The weights are calculated to minimize the mean squared difference between the education of Gurkha and synthetic Gurkha ethnic groups across ages 22 to 44. Based on the mathematical algorithm provided by **?**, I choose W to minimize $(X_G - X_N W)(X_G - X_N W)$, where $W = \{(w_1, ...w_J)'$ subject to $w_1 + ... + w_J = 1,\ w_J \geq 0\}$. X_G is a $(k\ X\ 1)$ vector of average years of education at each age ineligible birth cohorts for the Gurkha ethnic group, where $21 \leq k \leq 44$. X_N is a $(k\ X\ J)$ matrix with average years of education for k ineligible birth cohorts and J non-Gurkha ethnicities.

CHAPTER II

Sibling Rivalry in Education: Estimation of Intrahousehold Trade-offs in Human Capital Investment

2.1 Introduction

Households are key decision-makers of their children's human capital investment, which is a crucial determinant of their welfare and their country's economic growth. Sub-Saharan African countries spent on average 5% of its gross domestic product on education in 2008.[1] Despite this, half of all children still do not complete primary education in more than one-third of these countries. Improving educational outcomes for children requires a better understanding of the mechanisms that govern their households' investment decisions.

Standard household models developed by *Becker* (1981) and *Newman and Gertler* (1994), identify various factors that affect schooling, one of which is the trade-off among household members. The theoretical model suggests that a child's educational outcome could affect the education of other children living in the same household. An increase in the child's education shifts resources away from other children in

[1] The emphasis on education is illustrated by the UN's Millennium Development Goals (MDGs), where two of the eight goals focus on education–achieve universal primary education and achieve gender equality at all levels of education by 2015.

the household. It decreases the household's current income by limiting the child's ability to participate in the labor market. While these two mechanisms suggest a negative trade-off between siblings' education, it could also have a positive impact by decreasing the per-child fixed cost of education, increasing the permanent income of the household, and decreasing the marginal benefit from leisure for other children. Therefore, an increase in the child's education can either raise or lower the education of other children in the household and estimating whether this effect is positive or negative becomes an empirical question.

The causal estimation of the "inter-sibling" effect in education is, however, constrained by endogeneity concerns. *Strauss and Thomas* (1995) point out that there are significant unobserved heterogeneity across households that affect their members' educational outcomes. Hence using the child's education to estimate its impact on the education of other children in the household leads to a biased estimate because unobserved household characteristics are likely to be correlated with the education of all the children in the household. Moreover, the child's education affects other children's education and vice versa. *Manski* (1993) argues that this reflection problem could lead to a biased estimate of the effect of child's education on other children living in the household.

The ideal experiment requires two groups of households from developing countries– treatment and control households, that are similar in all socio-economic characteristics. Furthermore, the children in each households are randomly divided into two groups– targeted and non-targeted children. The targeted children from the treatment household experience an exogenous increase in educational attainment. The difference in education between the non-targeted children in the treatment household and the non-targeted children in the control household could be interpreted as the inter-sibling effect in education.

A unique event occurred in Nepal in 1993 that resembled key aspects of the ideal experiment. The British Gurkha tradition involves an annual recruitment of Nepali men by the British government into the British Armed Forces. Due to historical reasons, the majority of British Gurkha soldiers are comprised of 5 out of 15 ethnic groups of Nepal. These five ethnic groups are of Mongolian-Tibeto origin and are referred to as the "Gurkha" ethnic group for the remainder of this paper. All the other ethnic groups are referred to as "non-Gurkha"s.

The natural experiment involves a change in British Gurkha Army recruitment regulations. In 1993, the British government changed the educational requirement for the British Gurkha Army applicants, from requiring no education to requiring a minimum of 8^{th} grade education. The change was prompted by a broader ideological shift within the British Army regarding the importance of educated and well-trained soldiers[2]. As a result, the timing of the change is exogenous to social, economic, or political characteristics of Nepal and mainly affected the men from the Gurkha ethnic group who were age eligible to apply to the British Gurkha Army in 1993.

This creates a unique natural experiment in which age eligible men from the Gurkha ethnic group exclusively experienced an exogenous change in their educational returns and as a consequence, change in their educational outcomes. More importantly, since women are ineligible to join the British Gurkha Army, women of Gurkha ethnic groups were only affected by the rule change through an increase in the education of age-eligible Gurkha men living in their household. As a consequence, Gurkha females with age eligible male sibling comprise the treatment group; whereas, the control group includes the Gurkha females who do not have age eligible male siblings and all non-Gurkha females.

The results suggest that the existence of at least one age eligible male sibling in the household decreased the education of Gurkha girls by 0.11 years, which implies a decrease of 7% from their mean education of 1.56 years. Moreover, the results

[2] "King's Statement to the Commons on Future Defense Structure." The Times. July 26, 1990

indicate that the estimated negative inter-sibling effect decreases with greater child spacing, increases with greater number of age eligible male siblings, increases with household's involvement in agricultural production, and decreases with more remote access to school.

The rest of the paper is structured as follows: Section 2 discusses the household models that highlight the interaction among its members. Section 3 describes the natural experiment used for causal identification. Section 4 describes the data used for empirical estimation. Section 5 presents the empirical results. Section 6 presents robustness for the identification strategy. Section 7 identifies potential mechanisms that jointly determine the results and Section 8 concludes.

2.2 Previous Literature

The important role of education in fostering long-term economic growth in developing countries have motivated economists and policymakers to examine how households could be encouraged to increase their demand for their children's education. *Angrist and Lavy* (1999) use Maimonides rule to determine the class size in Israel and find that reducing class size had positive impact on test scores of fourth and fifth graders. *Kremer et al.* (2009) find that introducing merit based scholarships for girls in Kenya significantly increased their school attendance and academic exam scores. They also find positive spillover effects among girls who were unlikely to win the scholarship and boys who were ineligible for the scholarship. *Duflo and Hanna* (2005) find that increasing teacher attendance in India increased student grades and improved their graduation rate by 62%. *Schultz* (2004a) suggest that the Progresa program, a conditional educational grant, in rural Mexico increased education by 10%.

These studies examine the impact of educational policies but ignore the spillover effect of these policies on non-targeted household members. *Angelucci et al.* (2009)

find that the Progresa program only affected the schooling of secondary school-age children whose extended household included the primary school-age children receiving cash transfers. They identify the redistribution of resources within the family network as the main mechanism driving this result. This suggests that the interactions among household members and the division of resources among them are crucial in determining how household members respond to various education policies. It also suggests that education interventions could create spillover effects on non-targeted household members, which might lead to a over- or under-estimation of the overall impact of these policies.

The models of household behavior have been critical in identifying numerous mechanisms that govern the household decision making process. Household production models (*Becker*, 1965; *Gronau*, 1977; *Becker*, 1981; *Newman and Gertler*, 1994) suggest that households maximize their welfare given their resource constraints. One of the key aspects of these theoretical models is the allocation of time of household members across various market and non-market activities, including leisure, labor, home production, and human capital investment.

The division of labor and resources among household members is determined by the members' characteristics and market factors. According to *Becker* (1965, 1981), household members who are relatively more efficient at certain activities would also spend more of their time on those activities compared to other members. This suggests that the household members' socio-economic choices, which the household has control over, including human capital investment, are related to observed and unobserved household characteristics and therefore, should be treated as endogenous.

Bommier and Lambert (2000) find that the differences in male and female marriage markets in Tanzania could explain the gender based differences in school enrollment age. They find that households send their daughter to school at a younger age and also leave school at a younger age than their sons in order to maximize the present value of the bride price. *Behrman et al.* (1982) find that parents spend more household resources on the children with relatively smaller genetic endowment because they value their off-springs' earnings equality. These findings underscore the interdependency of a household member's socio-economic outcomes with the household preferences and household members' characteristics.

The household production model also highlights the trade-offs among household members in socio-economic activities, such as consumption, leisure, and human capital investment. *Becker* (1965) points out that an increase in the relative efficiency of any member on performing certain activities would result in a reallocation of the time spent by all other members away from those activities. *Newman and Gertler* (1994) find that the time worked on farm by any family member affects the marginal benefit to farm work for all the family members. They also find that an increase in the leisure of a member affects the marginal rate of substitution between leisure and consumption for other household members. These results suggest that the household members' choices are interrelated and the interactions among household members are crucial in determining household outcomes, including its members' human capital investment.

Pitt et al. (1990) investigate the effect of child morbidity on the activities of other household members. They find that an increase in child morbidity in a household does not affect male siblings but affects mother and female siblings, where the mother shifts out of labor force into home care and female siblings drop out of school into the home production. *Rosenzweig and Wolpin* (1980), *Lee* (2008), and *Qian* (2009) study the effect of an increase in the number of children on the household on child quality, including human capital investment. *Rosenzweig and Wolpin* (1980) use twin births as an exogenous increase in the number of children to find that it has a negative impact on schooling in rural India. *Lee* (2008) uses the son preferences and the first child's gender to determine that a greater number of siblings have adverse effects on per-child investment in education. Using the relaxation of one child policy in rural China as an exogenous increase in family size, *Qian* (2009), on the other hand, finds that an additional sibling increases school enrollment of the first child by 8-17%.

2.3 Empirical Strategy

The effect of increasing a child's education on the education choices of other children in the household is ambiguous. The educational outcome of the other children is partly governed by the intra-household interactions, triggered by the increase in education of the child, as suggested by the earlier studies (*Becker*, 1965; *Behrman et al.*, 1982; *Pitt et al.*, 1990; *Newman and Gertler*, 1994; *Lee*, 2008; *Angelucci et al.*, 2009; *Qian*, 2009). Some of the mechanisms, such as the redistribution of household resources and decrease in household income, suggest a negative impact of increasing a child's education on the education of other children; whereas other mechanisms, such as the lowering of per child fixed costs of schooling, increase in permanent household income, and the economies of scale in leisure, indicate a positive inter-sibling effect in education. Hence, an increase in the chid's education can either raise or lower the education of other children in the household and estimating whether this effect is positive or negative becomes an empirical question.

The causal estimation of the inter-sibling effect in education is, however, constrained by endogeneity concerns because of the unobserved heterogeneity across households that are correlated with the education of all their members as indicated by *Strauss and Thomas* (1995) and due to the reflection problem highlighted by *Manski* (1993). The ideal experiment requires dividing the households with multiple children into two random groups– treatment and control households. Furthermore, the children in only the treatment households are randomly divided into two groups– treatment and control children. The treatment children in these households are subject to an exogenous increase in education. Under these circumstances, the difference in the educational attainment between the control children in the treatment households and the children in the control households could be interpreted as the causal estimate of the inter-sibling effect in education.

A unique event occurred in Nepal in 1993 that resembled key aspects of the ideal experiment. It involves a change in the education requirement affecting the recruitment of Nepalese men by the British government into the British Armed Forces.

2.3.1 British Brigade of Gurkha

Nepal is a landlocked country surrounded by India on three sides and China to its north. Its geographical position historically made it a melting ground for people and cultures from both north and south of its border (*Shrestha*, 2001b). The 1996 National Living Standard Survey categorizes the population of Nepal into 15 ethnic groups. Out of them, the Gurkha ethnic group is comprised of 5 Mongoloid tribal groups—the Rai, Limbu, Gurung, Magar, and Tamang, who settled in the eastern and central hills of Nepal during the initial wave of migration from north.

The recruitment of Nepalese men into the British Army started after the Anglo-Gurkha war (1814-1817). The treaty of Sugauli following the end of the war in 1817 allowed the British to set up three Gurkha regiments in the British Indian Army. The early recruits included ethnic groups such as the Rajput, Thakuri, Chetri, and Brahman, who migrated from south and were closely associated with the ethnicities found in India. In 1857, Indian soldiers serving in the British Indian Army led a mutiny against the British rule. Although the rebellion was eventually contained in 1858, the British became wary of Indian nationals serving in their army. *Rathaur* (2001) and *Caplan* (1995) argue that as a result, the British also stopped recruiting Nepalese men belonging to the ethnicities that originated from India into the British Gurkha Army. According to *Rathaur* (2001) after 1857, the new Nepali recruits in the British Gurkha Army were mainly drawn from the Rai, Limbu, Gurung, Magar, and Tamang ethnic groups who, unlike other ethnicities in Nepal, had migrated from north. This ethnicity bias in the recruitment of the British Gurkha Army continues to exist till present day as majority of the current British Gurkha soldiers are comprised of these 5 Mongolian-Tibeto ethnic groups.

The benefits from joining the British Gurkha Army is substantial. The present value of the lifetime income of serving in the British Gurkha Army for 25 years is estimated to be about $1,522,750 (*Shrestha*, 2011). This includes a starting salary of $21,000 and a lifelong annual pension of about $15,000 after serving for 25 years. According to *Caplan* (1995), remittances from Gurkha soldiers and pensions for ex-Gurkha soldiers were the country's largest earner of foreign currency until recent development of tourism and other sources of migration. More importantly, the pay and pensions of the servicemen are the major source of capital in most Mongolian-Tibeto communities in the hills of Nepal whose main alternative employment is limited to farming.[3] The financial benefits of British Gurkha Army in these communities is evident from quotes documented by *Caplan* (1995) such as, "One of my boys has gone to the Army, we have only that hope."

2.3.2 Natural Experiment: A Change in Education Criteria

Education is an important aspect of the selection process. Since 1997, the recruits must have completed at least 10 years of education. However prior to 1993, no formal education was required to join the British Gurkha Army. Moreover, the selection process was strictly limited to physical examinations. In 1993, the British Army changed its educational requirement and the new Gurkha recruits were required to have completed at least 8 years of education. Following the change in educational requirement, mathematics and English grammar tests were added to the selection

[3] The Defense Committee of the British House of Commons in 1989 suggested that the annual salary of a British Gurkha soldier was about 100 times the average income in the hills from where they come from.

process.

The change in educational requirement was a result of a series of defense reviews, termed “Option for Change,” conducted by the UK Ministry of Defense in early 1990s. The Option for Change led to an extensive restructuring of the British Armed Forces in order to cut defense spending after the end of the Cold War. It led to a reduction in service manpower by 18% to an Army of around 120,000[4]. More importantly, it focused on maintaining a leaner, trained, and educated army with better technological capabilities[5].

Since the change in the British Gurkha education requirement was instigated by a broader restructuring of the British Army, the timing of the change is unlikely to be correlated to any social, economic, or political characteristics of Nepal at that time. As a result, this change created an exogenous increase in educational returns abroad for a subset of individuals in Nepal.

Shrestha (2011) finds that the change in the education requirement to join the British Gurkha Army led Gurkha men to increase their education by an average of 1.24 years. It induced 19.5 percent of Gurkha men who would not have received any formal education in the absence of the rule change, to enroll in school. *Shrestha* (2011) shows that the impact of the rule change was significant among Gurkha men of age 12 or younger in 1993, because they were likely to be enrolled in school at the time of the rule change and also had sufficient number of years to meet the new educational requirement. Individuals who were 12 or younger in 1993 are referred to as “age-eligible” for the remainder of this paper.

[4]The Eighth Defense Report published by the Defense Committee, House of Commons in 1997.

[5]The shift towards better-educated soldiers in 1990s was not unique to the British Army but was also prevalent in the US Army. According to the Tenth Quadrennial Review of US Military Compensation, the number of recruits with high school diploma increased by 30% in mid-1980s and those who scored better than the median in the Armed Forces Qualification Test (AFQT) increased by 10% in early 1990s

2.3.3 Identification Strategy

The empirical analysis seeks to measure the inter-sibling effect in education by estimating the impact of age-eligible Gurkha men increasing their level of education on Gurkha females living in the same household. Since the British Gurkha Army recruitment is limited to Gurkha men, the education of Gurkha females would only be affected by the rule change through an increase in the education of their male siblings. Hence, the exposure to the rule change among females is jointly determined by ethnicity and whether they have any age-eligible male sibling living in the household.

The effect is identified via difference-in-difference estimation, comparing the educational attainment between females who have age-eligible male sibling and who do not have age-eligible male sibling, across Gurkha and non-Gurkha ethnic groups. The identification assumption is that in the absence of the rule change in 1993, the difference in female education between Gurkha and non-Gurkha ethnic groups would have been identical across households with and without age-eligible men.

The simple comparison of education between Gurkha and non-Gurkha females with age-eligible male sibling is likely to be driven by unobserved ethnic characteristics. The diff-in-diff strategy controls for the unobserved ethnic characteristics that could affect education as well as controls for the unobserved household characteristics that could be correlated with fertility and education. Hence, the diff-in-diff estimate could be interpreted as the causal effect of an increase in the education of age-eligible Gurkha men, triggered by the exogenous change in the educational requirement to join the British Gurkha Army, on the education of Gurkha females living in the household.

The above identification strategy can be expressed using the following regression framework:

$$Y_{iklmh} = c + \alpha_{1k} + \beta_{1m} + \eta_{1l} + \delta\, T_{ih} + \gamma\, (G_{im} * T_{ih}) + \Sigma\, R_h + \Lambda\, K_i + \epsilon \quad (2.1)$$

where Y_{iklm} is the outcome of interest for woman i of ethnicity m in household h and age k; α_{1k} is an age dummy for each k; β_{1m} is an ethnicity dummy for each m; η_{1l} is a district of birth dummy for each l; G_{im} is a dummy indicating whether individual belongs to the Gurkha ethnic group; T_i is a dummy indicating whether the individual has any "age-eligible" male sibling; R_h is a vector of household-specific variables; and K_l is a vector of individual characteristics. The coefficient γ estimates the inter-sibling effect.

The diff-in-diff estimate could be biased if unobserved characteristics that are correlated with education and fertility differ systematically across Gurkha and non-Gurkha ethnic groups. *Shrestha* (2011) show that Gurkha men of only specific age, i.e. age-eligible Gurkha men, were affected by the rule change. Hence, the paper employs a diff-in-diff-in-diff strategy by utilizing the specific age of the male sibling as a third dimension that determines the exposure to the rule change. This involves subtracting the diff-in-diff estimate, calculated from Gurkha and non-Gurkha ethnic groups across females with no male siblings and females with any male siblings (age-eligible and age-ineligible), from the difference in education between Gurkha and non-Gurkha females with age-eligible male sibling.

The new identification strategy can be generalized in the following regression framework:

$$Y_{iklmh} = c + \alpha_{1k} + \beta_{1m} + \eta_{1l} + \delta\, T_{ih} + \theta\, P_{ih} + \sigma\,(G_{im} * P_{ih}) + \lambda\,(G_{im} * T_{ih}) + \omega\,(T_{ih} * P_{ih}) + \gamma\,(G_{im} * P_{ih} * T_{ih}) + \Sigma\, R_h + \Lambda\, K_i + \epsilon \qquad (2.2)$$

where P_{ih} is a dummy indicating whether the individual i has any male sibling of any age (age-eligible or age-ineligible) living in the household h.

The coefficient γ estimates the net effect of an increase in the child's education on the education of other children living in the household, which is jointly determined by numerous off-setting mechanisms. On one hand, an increase in the child's education shifts resources away from the education of other children in the household. It also limits the child's ability to participate in the labor market which decreases the current income of the household. On the other hand, it has a positive impact by decreasing the per child fixed cost of education, by increasing the permanent income of the household, and by decreasing the marginal benefit of leisure for other children.

2.4 Data

The paper uses the data from 2001 Census of Nepal. The empirical analysis focuses on females who were 6 years old to 12 years old in 2001. The census data are supplemented with 1996 Nepal Living Standards Surveys (NLSS), which is a household sample survey with greater detail. Table 2.1 presents summary statistics for the 235,297 girls from the 2001 Census as well as the 3,373 households from the 1996 NLSS.

Individual and household level averages for some socio-economic characteristics are provided for the entire sample as well as separately for the Gurkha and non-Gurkha ethnic groups. The Gurkha ethnic group comprises of just under 20 percent of the samples in both surveys. Panel A shows that the average level of education for the Gurkha ethnic group is 1.53, which is slightly lower than the average for the non-Gurkha ethnic group of 1.59. Similarly, 79.3% of Gurkha girls were born in rural area compared to 64% of non-Gurkha girls. 65% of Gurkha girls have age-eligible male sibling compared to 65.7% of non-Gurkha girls. The average household size for Gurkha and no-Gurkha ethnic groups are 6.16 and 6.30 respectively. Panel B suggests that near the time of the rule change, non-Gurkha households had better access to school facilities, health clinics, and transportation than Gurkha households. 48.5% of Gurkha households were living in poverty compared to 46.9% of non-Gurkha households.

2.5 Results

The diff-in-diff strategy compares the educational attainment of females who have at least one age-eligible male sibling in the household to those who do not have any age-eligible male sibling, in both Gurkha and non-Gurkha ethnic groups. Table 2.2 shows average years of education completed as of 2001 for females of age 6 to 12 in 2001.The empirical analysis focuses on girls 6 to 12 years old for three main reasons. First, they are of similar age compared to men who were affected by the rule change, and as a result are most likely to complete with age-eligible men for household resources including human capital investment. Second, they are old enough to have the opportunity to enroll in school by 2001 and have data on education. Third, they are young enough to be not married in 2001 and hence, have information on male siblings who grew up with them[6].

The results in table 2.2 show that an average education for Gurkha girls with age-eligible male sibling is 1.62 years, which is less compared to non-Gurkha girls with age-eligible male sibling whose average education is 1.72 years. This difference of 0.09 years is statistically significant at the 5 percent level. On the other hand, among girls with no age-eligible male sibling[7], there is no statistical difference in the average education across Gurkha and non-Gurkha ethnic groups. This suggests that in the absence of the rule change, the female educational attainment would not have differed across Gurkha and non-Gurkha ethnic groups, even among girls with age-eligible male siblings. The simple difference-in-differences estimation shows that girls

[6] The information on male siblings is obtained from the household roster. As a result, if a women gets married and moves into a new household with her spouse before 2001, then the household roster only provides information of the members living with her in 2001 but not the members who she grew up with before she got married. Among girls of age 6 to 12 in 2001, 0.7% are married and as a result, are dropped from the empirical analysis.

[7] The sample of girls with no age-eligible male sibling includes females who have no male sibling or who have only age-ineligible male siblings.

of Gurkha ethnic group with at least one age-eligible male sibling completed 0.11 less years of education. This is significantly different from zero at the 1% level.

Table 2.3 presents the difference-in-difference strategy by estimating coefficient γ in equation (2.1) for dependent variable– years of education completed as of 2001. The specification in column 1 does not control for any individual and household characteristics and hence, the diff-in-diff estimate is identical to the estimate of table 2.2. The specification in column 2, on the other hand, controls for cohort of birth dummies and ethnicity dummies because the difference in age distribution and ethnicity composition of girls with and without age-eligible male siblings, might not be consistent across Gurkha and non-Gurkha ethnic groups. The estimates in column 2 suggest that education of Gurkha girls with age-eligible male siblings decreased by 0.09 years, which implies a decrease of 6% from a mean education of 1.56 years. The estimate is statistically significant at the 1% level.

Figure 2.1 shows that Gurkha ethnic group is concentrated in the mid-north and north-east districts of Nepal. Table 2.1, panel A suggests that girls from Gurkha ethnic groups are less likely to be born in an urban region than non-Gurkha girls. Since the regional characteristics could be correlated with fertility and education, the specification in columns 3, 4, 5, and 6 includes district of birth dummies and a dummy indicating whether a girl was born in a rural region. The estimate in column 3 do not change and remains statistically significant at the 1% level.

Table 2.1 also suggests that among girls who have at least one age-eligible male sibling, Gurkha girls are likely to have slightly more number of age-eligible male siblings than non-Gurkha girls. The unobserved characteristics that are correlated with fertility and education could be different across Gurkha and non-Gurkha ethnic groups. The specification in columns 4, 5, and 6 controls for the household size, the number of age-eligible male siblings, and the number of male siblings of any age; whereas the specification in columns 5 and 6 additionally controls for age spacing between the girl and her age-eligible male sibling as well as age gap between her and her male sibling of any age. The results in columns 4 and 5 suggest that having any age-eligible men in the household decreased Gurkha girls education by 0.09 years. This implies that an increase in educational returns for age-eligible Gurkha men reduced Gurkha girls education by 6%, from an average education of 1.56 years of completed schooling in 2001.

The specification in column 6 corrects for other ethnic characteristics, such as access to school, access to healthcare, and poverty. These ethnic-level characteristics are obtained from 1996 NLSS. The estimates in column 6 are also statistically significant at the 1% level and suggest that an exogenous increase in the the educational returns of age-eligible Gurkha men decreased education of Gurkha girls living in the same household by 0.08 years.

The above interpretation is based on the assumption in the absence of the rule change, the difference in education between Gurkha and non-Gurkha girls would have been identical across households with and without age-eligible men. The results in table 2.1, however, suggest that unobserved ethnic characteristics could be jointly correlated with fertility and education. Table 2.4 corrects for these unobserved ethnic characteristics by employing a diff-in-diff-in-diff strategy and estimating coefficient γ in equation (2.2).

Table 2.4 presents the estimates for two dependent variables and 6 different specifications similar to table 2.3. The estimates in all the columns of table 2.4 do not change considerably compared to the estimates in table 2.3 and are statistically significant at the conventional levels. The results in column 5 suggest that an increase in educational returns for age-eligible Gurkha men decreased the education of Gurkha girls with age-eligible male sibling by 0.11 years. The results indicate that an inter-sibling effect of human capital investment is likely to be negative.

2.6 Robustness

The validity of the empirical strategy relies on the assumption that fertility is unaffected by the change in educational requirement. Since fertility is measured through household roster, the interpretation of the above results as an inter-sibling effect of education requires that the change in the educational requirement of the British Gurkha Army in 1993 did not affect the number of individuals, especially male siblings, living together in the household among Gurkha ethnic group.

The change in the British Gurkha Army education requirement in 1993 could induce age-eligible or age-ineligible Gurkha men to leave the household. The data on number of male siblings living separately from the household is available for a subset of girls used in the empirical analysis. The subsample for whom the information regarding the number of male siblings living separately is available, includes 97.5% of Gurkha girls and 96.2% of non-Gurkha girls from the original sample.

Table 2.5 and table 2.6 estimates coefficient γ in equation (2.1) and (2.2) respectively for this subsample. In both tables, specification in columns 2, 5, and 6 includes a dummy indicating whether the girl has a male sibling living separately from the household. The results in table 2.5 and 2.6 indicate that controlling for male siblings who are not part of the household in 2001 does not change the estimates in any of the specifications. More importantly, the estimates in columns 2, 4, and 6 in table 2.5 and 2.6 are all statistically significant at the conventional level. This suggests that the results in table 2.3 and 2.4 can not be explained thought the changes in household structure of Gurkha ethnic group, mainly the departure of male siblings from the household triggered by the rule change in 1993, but are more likely to be driven by the intra-household sibling interactions induced by the increase in educational attainment among age-eligible male siblings living in the household.

The departure of male siblings from the household also implies that they are not counted as the girl's male siblings in the empirical analysis. If unobserved ethnic characteristics are correlated with education and the likelihood of having all age-eligible male siblings living separately, then the above results are likely to be biased. Table 2.7 compares the household composition between Gurkha and non-Gurkha ethnic groups. The results in table 2.7 show that both Gurkha and non-Gurkha girls are equally likely to have at least one age-eligible male sibling living in the household, suggesting the absence of differential household formation across the two ethnic groups. Moreover, figure 2.2 suggests that age-eligible male siblings, who are 20 years old or younger in 2001, are too young to live separately from their parents. According to figure 2.2, only 1.94% of Gurkha households are headed by men of age 20 or younger and 2.24% of non-Gurkha households are headed by age-eligible men.

Table 2.8 employs an alternate strategy to assess household formation bias. Columns 1 and 2 presents the triple difference estimates from table 2.6, columns 1 and 3 respectively. The specification in columns 3 and 4 assumes that all the male siblings living separately in 2001 belong to the age-ineligible cohort and more importantly, they would have lived in the same household as the girl in the absence of the rule change. The results in columns 3 and 4 suggest that including male siblings living separately as age-ineligible male siblings living in the household for all girls in the subsample, does not change the estimates compared to columns 1 and 2. Similarly, the specification in columns 5 and 6 assumes that all the male siblings living separately in 2001 belong to the age-eligible cohort and would have lived in the same household as the girl if not for the rule change. Under this assumption, the estimates in columns 5 and 6 of inter-sibling effect does not differ from the estimates in columns 1 and 2 and are also statistically significant at the 1% level. This suggests that the negative results in table 2.3 and 2.4 are not likely to be driven by differential attrition of male siblings from the household.

Finally, the paper employs a placebo study to examine whether the results could be explained by unobserved ethnic characteristics that are correlated with education and the likelihood of having any age-eligible siblings. Table 2.9 estimates γ in equation (2.2) but instead of comparing educational attainment of girls based on whether have have any age-eligible male siblings, it compares girls' education based on whether they have any age-eligible female siblings. Since age-eligible female siblings were not affected by the change in education requirement of the British Gurkha Army, the estimate of inter-sibling effect should be zero. The results in table 2.9 show that all of the estimates are statistically insignificant at the 5% level. Although the high rate of marriage attrition among age-eligible female siblings could biased these estimates, the results do not refute the validity of the identification assumption and further supports the claim that the results in table 2.3 and 2.4 are driven by the increase in education of age-eligible Gurkha men living in the household.

2.7 Mechanisms

The above results estimate the net effect of increasing age-eligible Gurkha men's education on the education of girls living in the household, which is jointly governed by numerous off-setting intra-household mechanisms. Some negative mechanisms identified by the theoretical models include the shifting of household resources away from the human capital investment of girls and the decrease in current household income due to a decease in labor supply of age-eligible Gurkha men. The positive mechanisms involve lowering the girl's share of the fixed costs of education, increasing the permanent household income, and decreasing the girl's marginal benefit of leisure.

Table 2.10 seeks to identify the increased competition of household resources among its members due to an increase in age-eligible Gurkha men's education. Standard household models suggest that an increase in educational returns for age-eligible Gurkha men leads to a greater investment in their education, using resources that could have been previously devoted to education of Gurkha girls in the household. These models suggest that the competition among resources, including human capital investment, is decreasing with greater child spacing. Table 2.10 estimates γ in equation (2.2) for Gurkha girls separately across age difference between them and their age-eligible male siblings.

The results in table 2.10 suggest that the estimates of the inter-sibling effect become less negative as the child spacing between the Gurkha girl and her age-eligible male sibling increases. According to column 3, the education of Gurkha girls with age-eligible male sibling who is not more than 1 year older or 1 year younger than the girl, decreased by 0.19 years and it is statistically significant at the 1% level. On the other hand, the education of Gurkha girls with age-eligible male sibling who is at least 8 years older than the girl, decreased by only 0.05 years and it is not statistically significant even at the 10% level. The F test in column 3 shows that the estimates of the inter-sibling effect across different child spacings are statistically different from one another at the 5% level.

Table 2.11 identifies the inter-sibling effect for Gurkha girls across different number of age-eligible Gurkha male siblings. The results from column 3 show that the education for Gurkha girls with exactly one age-eligible male sibling decreased by 0.10 years, whereas, the education for Gurkha girls with more than 4 age-eligible male siblings in the household decreased by 0.35 years. More importantly, the estimated negative effect of the inter-sibling interaction is more severe as the number of age-eligible male siblings in the household increases and the F test in column 3 suggests that these estimates are statistically different from one another at the 1% level. The result in table 2.11 could be explained partly through the greater resource competition among household members with larger number of age-eligible male siblings and partly because of larger decrease in household income as more number of age-eligible male siblings shift out of household enterprise towards schooling.

The effect of decreasing age-eligible male sibling's labor market participation on the household income and in turn, on the girl's education is further highlighted by the results in table 2.12. Table 2.12 examines the differential inter-sibling interaction in the education of Gurkha girls based on various socio-economic characteristics, including access to schooling and operation of household agricultural enterprise. Since most of the child labor are limited to household agricultural production, this implies that households that operate agricultural production experience a greater reduction in household income as age-eligible Gurkha men devote more of their time on schooling and away from household activities. The results in columns 6 and 7 indicate that education of Gurkha girls with at least one age-eligible male siblings living in the household that operates agricultural enterprise decreased by 0.12 years and statistically significant at the 10% level; whereas, for those living in the household that are not involved in agricultural production decreased by only 0.07 years, which is not statistically significant at the 10% level. The F stat, however, suggests that these two estimates are not statistically different from one another at the conventional level.

The results in table 2.12 also show that the estimated inter-sibling effect is less negative for households that are in the districts with remote access to school, measured in terms of average distance-time between the household and the school. Although the estimates in column 4 and 5 are not statistically different from one another at the 10% level, it does suggests that the ability to spread fixed cost of transport across multiple siblings (including girls) moderates the negative effect of having an age-eligible brother, but does not overturn it[8].

2.8 Conclusion

Education has been the focal point of development policies for the past few decades. Economist and policymakers alike have devoted significant resources on

[8]Although the fixed cost of schooling has multi-facets, the empirical analysis is focused on only one aspect of the fixed cost, i.e. time it takes to drop and pick up children to and from school. The districts with more remote access to school indicates a large fixed cost of sending children to school as parents have to devote more time transporting their children to school.

understanding and evaluating household's response to various educational policy interventions. Since households are the key decision makers of their children's educational outcomes, it is important to investigate these educational policies in the context of household decision-making process.

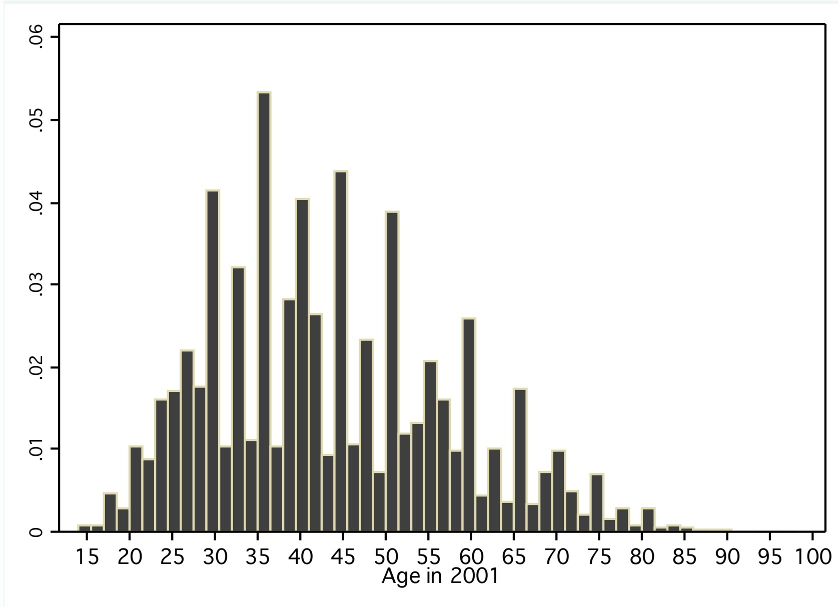
Although theoretical household behavior models highlight the interactions among household members regarding various socio-economic decisions, including human capital investment, the empirical estimation of how the education of one child affects the education of other children living in the household is absent due to endogeneity concerns. The paper utilized a unique natural experiment involving an increase in the educational requirement to join the British Gurkha Army for age eligible Gurkha men, to estimate the effect on the education of Gurkha girls living in the same household as them. Although the triple difference estimation suggests a robust negative inter-sibling effect, the result should be considered carefully given the unique nature of the British Gurkha recruitment, which is unlike any other labor markets or policy interventions.

The paper sought to identify some of the off-setting mechanisms that govern the inter-sibling interaction in education, including resource competition among its members, changes in household income, and sharing of the fixed cost of schooling. However, a more suitable natural experiment or a well-designed randomized control trial is required to precisely estimate the magnitudes of each of these mechanisms that jointly determine the inter-sibling effect in education.

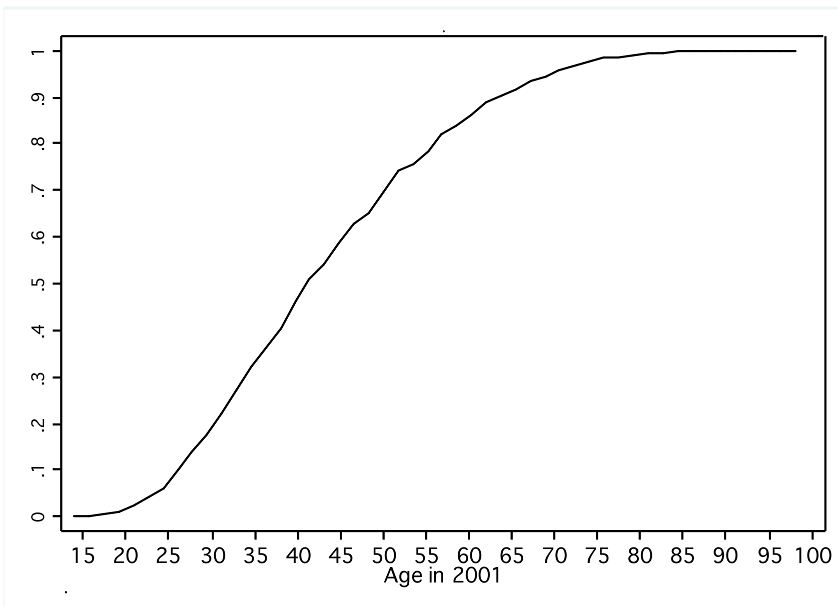
The paper highlights the importance of inter-sibling interactions in determining the socio-economic outcomes of its household members. It suggests a need to understand these interactions not only in educational outcomes but also other important variables such as health, nutrition, marriage, fertility, migration, and occupational choice. Analyzing these socio-economic decisions as a culmination of interactions between household members' preferences, choices, and resources, would not only help properly evaluate existing policy interventions but also help formulate more effective future policies.

Figure 2.1: Map of Nepal with Concentration of Gurkha Ethnic Population and Gurkha Recruitment Centers

Figure 2.2: Male Heads of Household by Age in 2001



(a) Probability Distribution Function



(b) Cumulative Distribution Function

Table 2.1: Descriptive Statistics

	Whole Sample	Gurkha	Non-Gurkha
Panel A: Individual Level Means			
2001 Census of Nepal			
Total Sample	235,297	44,250	191,047
% of sample	–	18.8%	81.2%
Level of Education	1.58	1.53	1.59
Percent Born in Rural	66.9%	79.3%	64.0%
Percent Married	0.7%	0.2%	0.8%
Age in 2001	8.97	8.99	8.97
Household Size	6.27	6.16	6.30
Number of Brothers	1.11	1.11	1.10
Number of Treated Brothers	1.00	1.01	1.00
Percent having Brother(s)	68.4%	67.6%	68.6%
Percent having Treated Brother(s)	65.6%	65.0%	65.7%
Percent Owning Land	75.5%	82.4%	73.8%
Percent Owning Livestock	72.0%	80.6%	70.0%
Panel B: Household Level Means			
1996 NLSS			
Total Sample	3373	544	2829
% of Sample	–	16.1%	83.9%
Access to School	0.38 Hrs	0.54 Hrs	0.35 Hrs
Access to Health Post	1.20 Hrs	1.85 Hrs	1.07 Hrs
Access to Paved Road	9.30 Hrs	14.45 Hrs	8.30 Hrs
Annual Consumption	$151.21	$149.93	$151.45
Percent of Household in Poverty	33.5%	48.5%	46.9%

Table 2.2: Means of Education by Ethnicity and Age-Eligible Brother

	Years of Education Completed[a]		
	Gurkha	Non-Gurkha	Difference
Age-Eligible Brother	1.62	1.72	-0.09
	(0 .010)	(0.005)	(0.012)
No Age-Eligible Brother	1.38	1.37	0.02
	(0.012)	(0.006)	(0.014)
Difference	0.24	0.35	-0.11
	(0.016)	(0.008)	(0.019)

Notes: Sample includes females aged 6 to 12 in 2001. Standard errors are reported in parentheses.

[a] *Highest level of education completed as of 2001 based on 2001 Census of Nepal*

* indicates significance at the 10 percent level

** indicates significance at the 5 percent level

Table 2.3: Effect of the Rule Change on Human Capital Investment of Gurkha Girls with Age-Eligible Brother (Difference-in-Difference)

Dependent Variable	(1)	(2)	(3)	(4)	(5)	(6)
Years of Education	-0.11**	-0.09**	-0.09**	-0.09**	-0.09**	-0.08**
	(0.053)	(0.035)	(0.040)	(0.039)	(0.040)	(0.037)
No. of Obs	231,777	231,777	229,915	229,915	229,915	225,388
Control Variables:						
Age Fixed Effects	No	Yes	Yes	Yes	Yes	Yes
Ethnicity Fixed Effects	No	Yes	Yes	Yes	Yes	Yes
District of Birth Fixed Effects	No	No	Yes	Yes	Yes	Yes
Rural	No	No	Yes	Yes	Yes	Yes
Household Size	No	No	No	Yes	Yes	Yes
No. of Brothers[a]	No	No	No	Yes	Yes	Yes
No. of Age-Eligible Brothers[b]	No	No	No	Yes	Yes	Yes
Brother Age Diff Dummies[c]	No	No	No	No	Yes	Yes
Age-Eligible Brother Age Diff Dummies[d]	No	No	No	No	Yes	Yes
Access to School[e]	No	No	No	No	No	Yes
Access to Health[f]	No	No	No	No	No	Yes
Poverty[g]	No	No	No	No	No	Yes

Notes: Sample includes females of age 6 to 12 in 2001. Standard errors are reported in parentheses. Standard errors are adjusted for within-caste correlation.

[a] ***No. of Brothers*** *is a variable indicating the number of brothers living in the household.*

[b] ***No. of Age-Eligible Brothers*** *is a variable indicating the number of treated males living in the household.*

[c] ***Brother Age Diff Dummies*** *are dummies indicating the age difference between the individual and her closest brother.*

[d] ***Age Eligible Brother Age Diff Dummies*** *are dummies indicating the age difference between the individual and her closest treated brother.*

[e] ***Access to School*** *is district-level average distance-time to school in 1996 calculated using 1996 NLSS*

[f] ***Access to Health*** *is district-level average distance-time to health post in 1996 calculated using 1996 NLSS*

[g] ***Poverty*** *is district-level average household consumption in 1996 calculated using 1996 NLSS*

* indicates significance at the 10 percent level

** indicates significance at the 5 percent level

Table 2.4: Effect of the Rule Change on Human Capital Investment of Gurkha Girls with Age-Eligible Brother (Difference-in-Difference-in-Difference)

Dependent Variable	(1)	(2)	(3)	(4)	(5)	(6)
Years of Education	-0.23**	-0.14**	-0.11*	-0.11*	-0.11*	-0.12**
	(0.080)	(0.063)	(0.060)	(0.061)	(0.060)	(0.062)
No. of Obs	231,777	231,777	229,915	229,915	229,915	225,388
Control Variables:						
Age Fixed Effects	No	Yes	Yes	Yes	Yes	Yes
Ethnicity Fixed Effects	No	Yes	Yes	Yes	Yes	Yes
District of Birth Fixed Effects	No	No	Yes	Yes	Yes	Yes
Rural	No	No	Yes	Yes	Yes	Yes
Household Size	No	No	No	Yes	Yes	Yes
No. of Brothers[a]	No	No	No	Yes	Yes	Yes
No. of Age-Eligible Brothers[b]	No	No	No	Yes	Yes	Yes
Brother Age Diff Dummies[c]	No	No	No	No	Yes	Yes
Age-Eligible Brother Age Diff Dummies[d]	No	No	No	No	Yes	Yes
Access to School[e]	No	No	No	No	No	Yes
Access to Health[f]	No	No	No	No	No	Yes
Poverty[g]	No	No	No	No	No	Yes

Notes: Sample includes females of age 6 to 12 in 2001. Standard errors are reported in parentheses. Standard errors are adjusted for within-caste correlation.

[a] ***No. of Brothers*** *is a variable indicating the number of brothers living in the household.*

[b] ***No. of Age-Eligible Brothers*** *is a variable indicating the number of treated males living in the household.*

[c] ***Brother Age Diff Dummies*** *are dummies indicating the age difference between the individual and her closest brother.*

[d] ***Age Eligible Brother Age Diff Dummies*** *are dummies indicating the age difference between the individual and her closest treated brother.*

[e] ***Access to School*** *is district-level average distance-time to school in 1996 calculated using 1996 NLSS*

[f] ***Access to Health*** *is district-level average distance-time to health post in 1996 calculated using 1996 NLSS*

[g] ***Poverty*** *is district-level average household consumption in 1996 calculated using 1996 NLSS*

* indicates significance at the 10 percent level

** indicates significance at the 5 percent level

Table 2.5: Effect of the Rule Change on Human Capital Investment of Gurkha Girls with Age-Eligible Brother Controlling for Separated Brother (Difference-in-Difference)

Dependent Variable	(1)	(2)	(3)	(4)	(5)	(6)
Years of Education	-0.09**	-0.09**	-0.09**	-0.09**	-0.09**	-0.09**
Obs: 226,075	(0.040)	(0.040)	(0.039)	(0.039)	(0.038)	(0.038)
Control Variables[a]:						
Separate Brother [b]	No	Yes	No	Yes	No	Yes
Household Size	No	No	Yes	Yes	Yes	Yes
No. of Brothers[c]	No	No	Yes	Yes	Yes	Yes
No. of Age-Eligible Brothers[d]	No	No	Yes	Yes	Yes	Yes
Brother Age Diff Dummies[c]	No	No	No	No	Yes	Yes
Age-Eligible Brother Age Diff Dummies[d]	No	No	No	No	Yes	Yes

Notes: Sample includes females of age 6 to 12 in 2001 for whom data on number of brothers living separately from the household are available. The data is available for females whose parents or grandparents are heads of the household. It comprises of 96% of the sample. Standard errors are reported in parentheses. Standard errors are adjusted for within-correlation.

[a] *All specifications include age fixed effects, ethnicity fixed effects, district of birth fixed effects, and rural dummy.*

[b] ***Separate Brother** is a dummy indicating whether the individual has brother(s) living separately.*

[c] ***No. of Brothers** is a variable indicating the number of brothers living in the household.*

[d] ***No. of Age-Eligible Brothers** is a variable indicating the number of treated males living in the household.*

* indicates significance at the 10 percent level

** indicates significance at the 5 percent level

Table 2.6: Effect of the Rule Change on Human Capital Investment of Gurkha Girls with Age-Eligible Brother Controlling for Separated Brother (Difference-in-Difference-in-Difference)

Dependent Variable	(1)	(2)	(3)	(4)	(5)	(6)
Years of Education	-0.21**	-0.21**	-0.21**	-0.21**	-0.20**	-0.20**
Obs: 226,075	(0.066)	(0.066)	(0.066)	(0.066)	(0.067)	(0.067)
Control Variables[a]:						
Separate Brother [b]	No	Yes	No	Yes	No	Yes
Household Size	No	No	Yes	Yes	Yes	Yes
No. of Brothers[c]	No	No	Yes	Yes	Yes	Yes
No. of Age-Eligible Brothers[d]	No	No	Yes	Yes	Yes	Yes
Brother Age Diff Dummies[c]	No	No	No	No	Yes	Yes
Age-Eligible Brother Age Diff Dummies[d]	No	No	No	No	Yes	Yes

Notes: Sample includes females of age 6 to 12 in 2001 for whom data on number of brothers living separately from the household are available. The data is available for females whose parents or grandparents are heads of the household. It comprises of 96% of the sample. Standard errors are reported in parentheses. Standard errors are adjusted for within caste-correlation.

[a] *All specifications include age fixed effects, ethnicity fixed effects, district of birth fixed effects, and rural dummy.*

[b] ***Separate Brother** is a dummy indicating whether the individual has brother(s) living separately.*

[c] ***No. of Brothers** is a variable indicating the number of brothers living in the household.*

[d] ***No. of Age-Eligible Brothers** is a variable indicating the number of treated males living in the household.*

* indicates significance at the 10 percent level

** indicates significance at the 5 percent level

Table 2.7: Effect of the Rule Change on the Probability of Having Age-Eligible Brother (Using Linear Probability Model)

Dependent Variable	(1)	(2)	(3)	(4)	(5)
Dummy Indicating	-0.00	-0.01	0.00	-0.00	-0.00
Age-Eligible Brother(s)	(0.011)	(0.012)	(0.009)	(0.008)	(0.009)
No. of Obs	231,777	231,777	231,777	231,777	227,213
Control Variables:					
Age Fixed Effects	No	Yes	Yes	Yes	Yes
District of Birth Fixed Effects	No	No	Yes	Yes	Yes
Rural	No	No	No	Yes	Yes
Access to School	No	No	No	No	Yes
Access to Health	No	No	No	No	Yes
Poverty	No	No	No	No	Yes

Notes: Sample includes females of age 6 to 12 in 2001. Standard errors are reported in parentheses. Standard errors are adjusted for within-caste correlation.
* indicates significance at the 10 percent level
** indicates significance at the 5 percent level

Table 2.8: Effect of the Rule Change on Human Capital Investment of Gurkha Girls with Age-Eligible Brother (Lee Bounds)

	Baseline		All Separate Brothers Age Ineligible[a]		All Separate Brothers Age-Eligible[b]	
Dependent Variable	(1)	(2)	(3)	(4)	(5)	(6)
Years of Education	-0.21**	-0.21**	-0.20**	-0.22**	-0.21**	-0.22**
	(0.066)	(0.066)	(0.063)	(0.063)	(0.065)	(0.065)
Control Variables[c]:						
Household Size	No	Yes	No	Yes	No	Yes
No. of Brothers[d]	No	Yes	No	Yes	No	Yes
No. of Elig Brothers[e]	No	Yes	No	Yes	No	Yes

Notes: Sample includes females of age 6 to 12 in 2001 for whom data on number of brothers living separately from the household are available. The data is available for females whose parents or grandparents are heads of the household. It comprises of 96% of the sample. Standard errors are reported in parentheses. Standard errors are adjusted for within-caste correlation.

[c] *All specifications include age fixed effects, ethnicity fixed effects, district of birth fixed effects, and rural dummy.*

[d] ***No. of Brothers*** *is a variable indicating the number of brothers living in the household.*

[e] ***No. of Age-Eligible Brothers*** *is a variable indicating the number of treated males living in the household.*

* indicates significance at the 10 percent level

** indicates significance at the 5 percent level

Table 2.9: Effect of the Rule Change on Human Capital Investment of Gurkha Girls with Age-Eligible Sister (Falsification Test)

Dependent Variable	(1)	(2)	(3)	(4)	(5)	(6)
Years of Education	0.14	0.15	0.10	0.10	0.21	0.13
	(0.145)	(0.104)	(0.085)	(0.086)	(0.138)	(0.136)
No. of Obs	231,777	231,777	229,915	229,915	227,515	223,018
Control Variables:						
Age Fixed Effects	No	Yes	Yes	Yes	Yes	Yes
Ethnicity Fixed Effects	No	Yes	Yes	Yes	Yes	Yes
District of Birth Fixed Effects	No	No	Yes	Yes	Yes	Yes
Rural	No	No	Yes	Yes	Yes	Yes
Household Size	No	No	No	Yes	Yes	Yes
No. of Sisters[a]	No	No	No	Yes	Yes	Yes
No. of Age-Eligible Sisters[b]	No	No	No	Yes	Yes	Yes
Sister Age Diff Dummies[c]	No	No	No	No	Yes	Yes
Age-Eligible Sister Age Diff Dummies[d]	No	No	No	No	Yes	Yes
Access to School[e]	No	No	No	No	No	Yes
Access to Health[f]	No	No	No	No	No	Yes
Poverty[g]	No	No	No	No	No	Yes

Notes: Sample includes females of age 6 to 12 in 2001. Standard errors are reported in parentheses. Standard errors are adjusted for within-caste correlation.

[a] ***No. of Sisters*** *is a variable indicating the number of female siblings living in the household.*

[b] ***No. of Age-Eligible Sisters*** *is a variable indicating the number of age-eligible (12 or younger in 1993) female siblings living in the household.*

[c] ***Sister Age Diff Dummies*** *are dummies indicating the age difference between the individual and her closest female sibling.*

[d] ***Age Eligible Sister Age Diff Dummies*** *are dummies indicating the age difference between the individual and her closest age-eligible female sibling.*

[e] ***Access to School*** *is district-level average distance-time to school in 1996 calculated using 1996 NLSS*

[f] ***Access to Health*** *is district-level average distance-time to health post in 1996 calculated using 1996 NLSS*

[g] ***Poverty*** *is district-level average household consumption in 1996 calculated using 1996 NLSS*

* indicates significance at the 10 percent level

** indicates significance at the 5 percent level

Table 2.10: Differential Effect of the Rule Change on Human Capital Investment of Gurkha Girls with Age-Eligible Brother (By Age Difference)

	(1)	(2)	(3)	(4)	(5)	(6)
$0 \leq AgeDiff \leq 1$	-0.27**	-0.22**	-0.19**	-0.18**	-0.17**	-0.20**
	(0.083)	(0.075)	(0.075)	(0.071)	(0.073)	(0.076)
$2 \leq AgeDiff \leq 3$	-0.24**	-0.15**	-0.12**	-0.11*	-0.11*	-0.12**
	(0.082)	(0.058)	(0.059)	(0.057)	(0.057)	(0.057)
$4 \leq AgeDiff \leq 5$	-0.23**	-0.13*	-0.11	-0.10	-0.10	-0.11
	(0.090)	(0.068)	(0.065)	(0.063)	(0.063)	(0.065)
$6 \leq AgeDiff \leq 7$	-0.25**	-0.15*	-0.13*	-0.12*	-0.12*	-0.14**
	(0.080)	(0.074)	(0.073)	(0.070)	(0.070)	(0.071)
$8 \leq AgeDiff \leq 14$	-0.11	-0.07	-0.05	-0.05	-0.05	-0.07
	(0.113)	(0.098)	(0.093)	(0.091)	(0.091)	(0.094)
Prob $\geq$ F(All Coefficients Equal)	0.412	0.004	0.004	0.003	0.003	0.003
Control Variables:						
Treated Brother Age Diff Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Age Fixed Effects	No	Yes	Yes	Yes	Yes	Yes
Ethnicity Fixed Effects	No	Yes	Yes	Yes	Yes	Yes
District of Birth Fixed Effects	No	No	Yes	Yes	Yes	Yes
Rural	No	No	No	Yes	Yes	Yes
Household Size	No	No	No	No	Yes	Yes
No. of Brothers	No	No	No	No	Yes	Yes
No. of Treated Brothers	No	No	No	No	Yes	Yes
Access to School	No	No	No	No	No	Yes
Access to Health	No	No	No	No	No	Yes
Poverty	No	No	No	No	No	Yes

Notes: Sample includes females of age 6 to 12 in 2001. The dependent variable is the years of education completed as of 2001. Standard errors are reported in parentheses. Standard errors are adjusted for within-caste correlation.
* indicates significance at the 10 percent level
** indicates significance at the 5 percent level

Table 2.11: Differential Effect of the Rule Change on Human Capital Investment of Gurkha Girls with Age-Eligible Brother (By Number of Age-Eligible Brothers)

	(1)	(2)	(3)	(4)	(5)	(6)
No. of Treated Brother = 1	-0.22**	-0.13*	-0.10	-0.09	-0.09	-0.10*
	(0.085)	(0.064)	(0.065)	(0.061)	(0.061)	(0.062)
No. of Treated Brother = 2	-0.25**	-0.15**	-0.13**	-0.12**	-0.12**	-0.14**
	(0.076)	(0.065)	(0.063)	(0.061)	(0.061)	(0.061)
No. of Treated Brother = 3	-0.16	-0.17*	-0.16**	-0.15*	-0.15**	-0.17**
	(0.105)	(0.091)	(0.080)	(0.076)	(0.076)	(0.079)
No. of Treated Brother $\geq$ 4	-0.28**	-0.36**	-0.35**	-0.32**	-0.32**	-0.34**
	(0.122)	(0.101)	(0.101)	(0.095)	(0.095)	(0.096)
Prob $\geq$ F(All Coefficients Equal)	0.105	0.006	0.010	0.009	0.008	0.007
Control Variables:						
No. of Treated Brother Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Age Fixed Effects	No	Yes	Yes	Yes	Yes	Yes
Ethnicity Fixed Effects	No	Yes	Yes	Yes	Yes	Yes
District of Birth Fixed Effects	No	No	Yes	Yes	Yes	Yes
Rural	No	No	No	Yes	Yes	Yes
Household Size	No	No	No	No	Yes	Yes
No. of Brothers	No	No	No	No	Yes	Yes
Access to School	No	No	No	No	No	Yes
Access to Health	No	No	No	No	No	Yes
Poverty	No	No	No	No	No	Yes

Notes: Sample includes females of age 6 to 12 in 2001. The dependent variable is the years of education completed as of 2001. Standard errors are reported in parentheses. Standard errors are adjusted for within-caste correlation.
* indicates significance at the 10 percent level
** indicates significance at the 5 percent level

Table 2.12: Differential Effect of the Rule Change on Human Capital Investment of Gurkha Girls with Age-Eligible Brother (By Density, Distance to School, and Operation of Household Enterprise)

Whole	Density[a]		Distance to School[b]		Household Enterprise[c]	
Sample	Rural	Urban	$\leq$ Median	$\geq$ Median	Ag Land	No Ag Land
(1)	(2)	(3)	(4)	(5)	(6)	(7)
-0.12**	-0.10	-0.21*	-0.15*	-0.07	-0.12*	-0.07
(0.055)	(0.064)	(0.109)	(0.083)	(0.075)	(0.062)	(0.122)
231,777	155,362	76,415	129,734	102,043	152,472	74,741

Notes: Sample includes females of age 6 to 12 in 2001. The dependent variable is the years of schooling completed as of 2001. Standard errors are reported in parentheses. Standard errors are adjusted for within-caste correlation. All specifications include age fixed effects, ethnicity fixed effects, district of birth fixed effects, rural dummy, and household size.

[a] ***Density:*** *Sample divided into two categories based on population density in the area of birth: individuals born in rural area and urban area.*

[b] ***Distance to School:*** *Sample divided into two categories based on average distance-time to school in 1996: individuals born in districts with average distance-time to school $\leq$ median district-level time and districts with average distance-time to school $\geq$ median district-level time. The median district-level distance-time to school is 0.36 hours.*

[c] ***Household Enterprise:*** *Sample divided into two categories based on ownership of agricultural land and livestocks: individuals in household with agricultural land and livestocks and household with no agricultural land and no livestocks.*

* indicates significance at the 10 percent level

** indicates significance at the 5 percent level

CHAPTER III

Access to the North-South Roads and Farm Profits in Rural Nepal

3.1 Introduction

Rural transportation is an important development strategy in low-income countries. It enhances economic gains for agrarian households by providing easier access to both input and output markets. Moreover, such gains are likely to have a substantial impact on poverty reduction because the poor are concentrated in the rural agricultural sector, where transportation infrastructure is most scarce.[1]

The World Bank in 2005 invested more than $8 billion US dollars on infrastructure in developing countries, which accounts for 35 percent of their total lending (World Bank, 2006). But measuring economic benefits of these large investments is difficult because of reverse causality. Although past economic growths are often always accompanied by rapid expansion of infrastructure, such strong correlations do not imply that investments on infrastructure will promote growth. On one hand, comparing regions with various degrees of infrastructure could overestimate its effect on economic development because development in and of itself also increases demand

[1] 50 percent of the poor in Nepal live in rural areas and are involved in agricultural production. Rural areas are often characterized by rugged terrain, sparse population, and isolation, thereby making effective transportation infrastructure for market integration even more crucial.

for infrastructure. On the other hand, given *Burgess and Jedwab* (2010) find that political appointees considerably influence road allocation across region, if political power and economic wealth are negatively correlated, this will bias the estimate of the effect of infrastructure towards zero. I overcome the endogeneity bias by utilizing a unique geographic feature that exogenously determines the placement of road in rural Nepal and estimate the impact on its agricultural economy.

Nepal is one of the poorest countries in the world with a dominant agriculture sector and a sparse road network due to an extremely rugged terrain. Based on 2001 Nepal census data, 44 percent of its population including 3.3 million poor live in the hills, out of which 75 percent of them rely on agriculture for their livelihoods. The rough terrain poses a serious challenge for road construction, which is reflected by one of the lowest road densities in the world.[2] Nevertheless, in the last four decades the government constructed 15,308 kilometers of roads by relying significantly on international donors, expanding from a meager length of 376 km in 1951.[3]

The main purpose of these roads is to connect district headquarters, which are historical fort towns that currently serve as administrative centers. The government's 20-Year Road Plan in 2001 adopted five main objectives, the top of which was "strengthening political and administrative linkages." *Shrestha* (2001a) points out that the economic consideration is always secondary to the administrative consideration in constructing rural roads. More importantly, because all mountains in the region span north-south, it is cheaper to construct a north-south road, which follows one of such mountain ranges, relative to constructing an east-west road that

[2]In 2007, the road density in Nepal was 121 km/100 km^2 or 0.8 km per 1,000 people. Only 43 percent of the rural population have access to all-season road and the average distance to the nearest transportation stop in rural area is 30 kilometers. 60 percent have access to roads within two-hours walking distance, while 23 percent are beyond a four-hour walk.

[3]According to the World Bank, Nepal's road network annually increased by 6.7% between 1995 and 2004 with the largest expansion occurring in roads classified as district or rural roads, which grew annually by 11% during this period. Based on the "Twenty Year Road Plan" published by the Ministry of Physical Planning and Work, Department of Roads, the public road expenditure accounted for 12 percent of the total national budget in 2001.

has to cross multiple mountains via bridges.[4] This unique geographic feature has resulted in a road network that includes a single east-west highway that stretches the entire country and multiple north-south roads, which connect to this highway at various locations. Each north-south road connects several district headquarters that are situated directly north of its junction with the east-west highway.

I use this geographic north-south corridors to construct an instrumental variable to deal with the endogeneity issue of where the road is constructed. In particular, I divide the country into 10 equal north-south blocks and connect all the district headquarters within each block with its nearest district headquarter by using a straight line.[5] Dividing the country into such north-south blocks ensures that the district headquarters are connected along the north-south direction rather than east-west direction, as suggested by the geographic feature of mountains. I also connect the district headquarter that is closest to the east-west highway within each block to that highway via a shortest distance, which results in 10 north-south lines connecting all the district headquarters in the hills to the east-west highway. Although 10 north-south blocks are based on the actual political and administrative division of Nepal,[6] I repeat the algorithm using 12 north-south blocks and show that the results are robust to my choice of the algorithm parameter. Lastly, I show that the distance to these straight lines is a strong predictor of household's proximity to road. Moreover, the construction of the straight lines are based on two factors, administrative integration and geographic feature that favors north-south linkages, both of which are likely to

[4]In 2001, the cost of constructing two-lane bridge in the hill region was estimated to be half a million rupees per meter. In the same year, the construction of bridges comprised of more than 30 percent of the total government funding on road development, and 15 percent of the total road investments based on the estimates from the Ministry of Finance. This illustrates that the construction of bridges constitutes a significant cost burden on any road project's total budget.

[5]As seen in figures 3.3 and 3.5, the following algorithm leads to some district headquarters being connected to multiple other district headquarters because the nearest district headquarter pairs are not symmetric i. e. the district headquarter that is nearest to another district headquarter does not always imply that the later district headquarter is also the nearest district headquarter of the former.

[6]Nepal is divided into 5 political and administrative regions, as shown in figure 3.4. The five pairs of these adjoining north-south blocks coincide well with these 5 administrative regions.

be independent of the socio-economic characteristics of households in the hinterlands.

My empirical method improves the identification strategies used by previous studies. While *Banerjee et al.* (2009) and *Atack et al.* (2009) construct a straight line connecting important historical cities to predict modern railway network, their instruments coincide with historical trade corridors and therefore could be correlated with unobserved characteristics that are associated with living close to a historical route. In contrast, I confine the linkage to a north-south direction making it less likely to be correlated with historical trade routes, which stretch in both an east-west and a north-south directions according to *Stiller* (1976).[7] *Jacoby and Minten* (2008) include transportation costs directly into a nonparametric model, while arguing that the differences in transportation costs are due to geographic terrain. The costs are significantly higher at the top of the mountain relative to the bottom. However, elevation could be correlated with land productivity as well as household's unobserved characteristics, especially if poor households are more likely to settle at higher elevations due to cheaper access to those lands. Apart from the physical feature of the terrain, my empirical strategy also relies on an administrative integration policy to explain the variation in access to road, and therefore, allows for appropriate controls of terrain quality such as elevation and land gradient.

Using cross-sectional household data from 2011 Nepal Living Standard Survey and geo-spatial data from MENRIS, I estimate a 2SLS model to calculate the effect of roads on household profit from agriculture as measured by the farm value. If farms behave like assets,[8] the improvement in farm's future profits due to lower transportation cost should be reflected in its value. This approach is used by *Arnott and Stiglitz* (1981) and *Jacoby* (2000) to circumvent the daunting task of calculating

[7]Based on *Stiller* (1976) and *Shrestha* (2001a), communications between district headquarters were historically carried out through foot trails that stretched east-west as well as north-south by connecting the district headquarters via a shortest distance, and they constituted the central nerve of the hill economy prior to the construction of roads.

[8]The standard asset-pricing model states that the present value of a plot is equal to the discounted sum of its future profits. This hypothesis is tested on rented farm plots later in the paper.

actual farm profits. Household surveys in low-income countries lack accurate information on farm profits, partly because many input purchases and sales of output are conducted informally through non-monetary transfers.

My results suggest that decreasing a household's travel time to the nearest road by one percent increases its farm value by 0.25 percent. The 2SLS estimates are statistically significant at the 5 percent level and are robust to the algorithm parameters used to construct the line. While earlier infrastructure studies like *Michaels* (2008) and *Donaldson* (2010) focus on price convergence from the view point of trade framework, I estimate household-specific economic gains from infrastructure investments. Because most of the world's poor inhabit rural areas and agriculture accounts for a significant fraction of a rural economy, the results directly contribute towards the government's goal to reduce poverty. *Mansuri et al.* (2012b) and *Mansuri et al.* (2012a) find that farm productivity and commercialization are two important determinants of escaping poverty in rural Pakistan and Uganda, both of which are likely to be directly affected by access to transportation (*Fafchamps and Hill*, 2005; *Gollin and Rogerson*, 2010). These results are consistent with my findings and together they highlight the benefits of rural transportation investments on agriculture, poverty alleviation, and food security.

The rest of the paper is structured as follows: Section 2 describes the geographic feature that influences road network in Nepal. Section 3 presents a theoretical model that illustrates the benefits of road projects on farm profits. Section 4 explains the identification used for causal estimation and Section 5 describes the data used for this strategy. Section 5 presents the empirical results and Section 7 concludes.

3.2 Background

Nepal is a landlocked country with isolated, localized economies referred to by *Shrestha and Jain* (1977) as "pocket economies". This fragmentation is a direct manifestation of its extremely rugged terrain, which has seriously undermined the development of adequate communication and transportation infrastructure (*Shrestha*, 2004). 77 percent of its land surface is covered by mountains and hills, with an exception of the Terrai plain in the south. The elevation in the hill region changes sharply from only 80 meters above sea level to more than 8800 meters within a stretch of just 100 kilometers, making road construction very difficult. As a result, road density is only 14 km/100 km^2 or 0.8 km per 1,000 people, which is one of the lowest in the world (*Meyer*, 2008). Figure 3.1 illustrates the rough landscape that covers almost the entire country.

Despite such rugged terrain, almost half of the population live in the hills where poverty incidence of 35 percent is the highest among all the regions (NLSS, 2004). According to *Upadhyaya* (2010), majority of these poor own small, fragmented landholdings and rely on subsistence farming for their livelihoods.[9] Moreover, most of the hill lands are steeply sloped, highly vulnerable to soil erosion, and impervious to irrigation, leading to an extremely low farm productivity (UNDP, 2002). These factors, together with a poor access to input and output markets due to sparse road network, severely restrict the economic gains from farming and worsen the food security problem in the region.[10] Therefore, the government has emphasized the construction of roads in the hill region as an important means to fight poverty.[11]

One of such infrastructure projects is the construction of the East West Highway (EWH), which began in the late 1950s but was only completed in the early 2000s. EWH spans the entire country along its southern border with India in the Terrai plain, with the total length of 1024 kilometers. It is the longest and the most impor-

[9]46 percent of the land owning households in the hills own less than 0.5 hectare of agricultural land.

[10]Based on the estimates by the International Fund for Agricultural Development (IFAD), more than 48 percent of children under 5 years of age are malnourished in the rural areas of Nepal.

[11]The Government of Nepal's Interim Three Year Plan (2007-2010) strongly emphasizes the role of roads in reducing poverty in rural areas, by improving rural access and prompting higher agricultural output as well as non-farm income in remote hill areas of the country.

tant highway in Nepal. According to *Shrestha* (2004), it is the backbone of Nepal's road network and plays a significant role in fostering social, economic, and political integration, by connecting all the major regional economies.

Although expanding the road network is an important government strategy for poverty reduction, its chief purpose is geared towards administrative integration by stressing the linkage between district headquarters. These are historical fort towns, which are scattered throughout the hill tops, and were converted into administrative centers under the current political system. According to *Shrestha* (2001a), the economic consideration have and still remains secondary to the administrative consideration in determining the construction of rural highways. The national transportation policy of the government states that "high priority shall be given to completing north-south roads connecting all 75 district headquarters to the main highway."

3.2.1 North-South Road

North-South (NS) roads are the most cost effective way of connecting district headquarters due to a geographic feature of the region, especially after the completion of EWH. All mountain ranges in the hill region span north-south followed by deep and narrow valleys or rivers on either sides. Consequently, a rural road is constructed by following one of such mountains from north to south until the road joins EWH in the Terrai plain. In contrast to constructing roads that stretch east-west, NS roads cross fewer mountain ranges and therefore require fewer number of bridges, which significantly reduces costs. Based on figures from the Finance Ministry, bridge construction accounts for more than 30 percent of the government's funding on road projects and 15 percent of its total road expenditure, which suggests that building a bridge constitutes a significant cost burden on a project's budget. Therefore, district headquarters are connected to the national road network via NS roads that all feed into EWH. The Twenty Year Road Plan published by the Ministry of Physical

Planning, emphasizes the construction of NS roads and highlights the government's reliance on NS roads to achieve its goal of administrative integration.

Transportation and communications between district headquarters were historically carried out through numerous foot trails that stretch east-west as well as north-south, connecting the district headquarters via a shortest distance (*Stiller*, 1976). Figure 3.2 shows all the foot trails that span across the country. According to *Shrestha* (2001a), the network of walking paths across the hills consisted of rope bridges, ferries, and other means of crossing between mountain ranges and prior to EWH, they constituted the central nerve of the hill economy. This implies that the recently constructed NS roads are likely to be uncorrelated with these foot trails that formed the trade link between district headquarters in the past. More importantly, the construction of NS roads in order to connect the district headquarters to EWH led to an exogenous increase in access to road for villages in north-south hinterlands relative to those in east-west hinterlands.

3.3 Theory

Farmers cultivate a single crop using two factors of production, x kg per hectare of chemical fertilizer and l hours per hectare of labor. Using a fixed technology, their crop yield y (kg per hectare) is given by the neoclassical production function $y = f(x, l)$. All farmers trade agricultural output and fertilizer in a competitive market centers located at the juncture of a rural road and the main highway.[12] The transportation of goods between farms and their nearest market center involves two components– head-loading goods between the farm and the road using human porters and trucking goods along the road, the former constituting the majority share of the

[12]As explained in the earlier section, the road network in Nepal comprises of EWH, which runs through the entire country, with various NS roads feeding into this main highway from the north at various locations along its length.

total transportation cost.[13]

From a given farm, it takes h hours to reach the nearest road junction and the portage cost of goods is t Rupees/ (kg hours). If k and p be the prices of fertilizers and outputs in the market center, respectively, then the effective purchase price of fertilizer is $\widetilde{k} = k + th$ Rupees/kg and the effective selling price of output is $\widetilde{p} = p - th$ Rupees/kg.[14] Let w be the wage rate of agricultural work. The per hectare land rent, r, can be defined as the maximal profit that can be earned from a hectare of land,

$$r\,(w,\,\widetilde{k},\widetilde{p}) \equiv max_{l,x}\,(\widetilde{p}y - wl - \widetilde{k}x) \tag{3.1}$$

Moreover, using the envelope theorem, the relationship between land rent and travel time to the nearest road is given by:

$$\frac{d\,r(w,\widetilde{k},\widetilde{p})}{d\,h} = -t(y+x) < 0 \tag{3.2}$$

The above equation implies that an improved access to roads for farmers increases their agricultural profit and therefore, raises the per hectare rent (r) of their plots. The negative rent gradient is equal to the total portage cost per hour per hectare.

3.4 Empirical Strategy

While the theoretical analysis above is framed in terms of land rents that capture the economic profits from cultivation, I use plot values in the empirical analysis

[13]This feature is likely to closely match the reality on the ground for most rural farmers in the hills, where almost six million people or 23% of the population still live more than four hours' walk away from the closest road. The markets at intermediate points between the farms and the road are an inessential complication because goods still need to be head-loaded from the road junction, the cost of which will be reflected on their prices at these intermediate market centers.

[14]The price of goods, k and p, are assumed not to vary across the market centers along the main highway (EWH). However, this assumption is relaxed in the empirical analysis with the inclusion of district (regional) fixed effects.

instead.[15] According to the standard asset-pricing model, the present value of a plot is equal to the discounted sum of its future profits, given by

$$log\ (V) \quad = \ log\,[\,r(w,\,\widetilde{k},\widetilde{p})\,] \quad - \ log\ (b) \tag{3.3}$$

where V is the present market value and b is the constant discount rate. Although the above equation provides a valuable link between observable land values and mostly unobservable land rents, the relationship may not be valid in Nepal because of land and credit markets distortions. If land is the sole form of collateral for loans, its price could capture its collateral value on top of the discounted future revenue stream from its rent (*Chalamwong and Feder*, 1988). Therefore, proving the validity of the asset pricing model requires estimating the following regression

$$log\ (V) \quad = \nu_0 \ + \ \nu_1\, log(r) \ + \ \varepsilon \tag{3.4}$$

and testing whether the coefficient estimate $v_1 = 1$. If this condition holds, then estimating the log-rent function of a plot is equivalent to estimating its log-value function, up to a constant.

3.4.1 Instrumental Variable

The empirical strategy to overcome selection bias is based on the unique geographic feature of Nepal. As described in the section above, all mountain ranges stretch north-south that results in a cheaper construction of roads that also run in the same direction. I use this feature to predict the placement of rural roads and implement an instrumental variable strategy to estimate the economic gains from better

[15] Nepal Living Standard Survey 2011, a nationally representative survey, collects detailed information on each plot owned by the household, its soil and irrigation characteristics, and its market value measured by the survey question, “If you wanted to buy/sell a plot exactly like this, how much would it cost/fetch you? ”

road accessibility for agrarian households living in the hills.

I take district headquarters that lie north of EWH and connect them using a simple algorithm. First, I divide the country into 10 north-south blocks based on equal longitudinal spacing. The five pairs of these adjoining blocks coincide well with the 5 development regions of Nepal,[16] with each block embodying 6 district headquarters on average. Within each block, I draw a straight line from each district headquarter to the nearest other district headquarter. Some district headquarters could be connected to multiple district headquarters because the nearest-headquarter pairs are not always symmetric.[17] Furthermore, restricting the connections to a north-south block ensures that the district headquarters are connected in a north-south direction rather than an east-west direction, which is an important feature of the instrumental variable strategy. Lastly, I connect the district headquarter within each block that is closest to EWH with EWH via a shortest straight line. The following algorithm results in 10 separate north-south linkages connecting the district headquarters to EWH. Figure 3.3 illustrates the 10 geographic blocks, the straight lines connecting the district headquarters, and EWH; whereas, figure 3.4 shows the five development regions. While 10 north-south blocks are based on the actual administrative division of Nepal, figure 3.5 shows the outcome of a similar exercise, but using 12 geographic blocks instead of 10 for alternate north-south linkages.

The straight line serves a proxy for NS road in the hill region. Whether the line drawn this way coincides well with actual road network can be tested, by estimating the correlation between the distance to the line and distance to road using the

[16]Nepal is divided into 5 development regions: Eastern, Central, Western, Mid-Western, and Far-Western regions, starting from east to west. Figure 3.4 maps these 5 development regions.

[17]Suppose there are three district headquarters A, B, and C. If A is the nearest district headquarter to B and vice versa, whereas the nearest district headquarter to C is also A, then district headquarter A will be connected to both district headquarters B and C by using the above algorithm.

following regression framework:

$$log\,(T_{id}) \quad = \; \alpha\; + \eta_d\; + \beta\; log\,(D_{id})\; + \varepsilon \qquad (3.5)$$

where T_{id} is the distance to the nearest road for rural household i living in district d, D_{id} is the distance to the nearest line, and η_d is district fixed effects. More importantly, D_{id} is influenced by two factors– administrative need to connect district headquarters and the geographic feature that favors north-south linkages. Both of them are likely to be independent of the socio-economic conditions of communities in the hinterlands, providing a plausibly exogenous variation in access to infrastructure for households in these regions.

The use of $log(D_{id})$ variable as an instrument to estimate the effect of road on farm profits requires satisfying the exclusion restriction. One potential concern in this strategy is that D_{id} could be picking household's proximity to the district headquarter itself, which in turn, directly affects farm profits and market land values. To address this, I control for the distance to the nearest district headquarter as well as the distance to EWH. *Banerjee et al.* (2009) point out that the distance to the straight line joining two cities in China violates the exclusion restriction because it also coincides with the historical trade routes between those cities. In contrast, the historical routes connecting any two cities in Nepal are equally likely to stretch east-west as well as north-south,[18] while the line is restricted to a north-south direction. Nevertheless I control for the distance to the nearest foot trails, which constitute the main channel through which all trades were carried out between two cities before the construction of highways as pointed out by *Stiller* (1976). Finally, D_{id} could also be correlated with the physical attributes of the agricultural land that directly affect its productivity and profits, I control for geographical characteristics such as distance to river, elevation, and land gradient as well as productivity variables such as suitability

[18]See figure 3.2 for the mapping of historical trade routes across the country.

for rice plantation (the main staple crop of the region) and irrigation. Therefore, the 2SLS model to estimate of the effect of road on farm profits is given by

$$\begin{aligned} log\,(T_{id}) \;=\; & \alpha^{'} + \eta_{1d} \;+ \beta^{'}\; log\,(D_{id}) \;+ \gamma^{'}\; log\,(Q_{id}) \;+ \delta^{'}\; log\,(R_{id}) + \theta^{'} log\,(M_{id}) + \\ & \zeta^{'}\; log\,(F_{id}) \;+ \xi^{'}\; E_{pid} \;+ \vartheta^{'}\; G_{pid} \;+ \psi^{'}\; log\,(A_{pid}) \;+ \Phi^{'}\; \bar{X}_{pid} \;+ \varepsilon^{'} \end{aligned} \tag{3.6}$$

$$\begin{aligned} log\,(V_{pid}) \;=\; & \alpha \;+ \eta_{1d} \;+ \beta\; log\,(T_{id}) \;+ \gamma\; log\,(Q_{id}) \;+ \delta\; log\,(R_{id}) \;+ \theta\; log\,(M_{id}) + \\ & \zeta\; log\,(F_{id}) \;+ \xi\; E_{pid} \;+ \vartheta\; G_{pid} \;+ \psi\; log\,(A_{pid}) \;+ \Phi\; \bar{X}_{pid} \;+ \varepsilon \end{aligned} \tag{3.7}$$

where V_{pid} is a market value of plot p owned by household i living in district d; η_{id} is a district dummy for each d; Q_{id} is the distance to the nearest district headquarter for household i in district d; R_{id} is the distance to the nearest river; M_{id} is the shortest distance to EWH; F_{id} is the distance to the nearest walking trails; E_{pid} is an elevation in meters; G_{pid} is a land gradient; A_{pid} is a total area of plot p owned by household i in hectares; $\bar{X}_{pid}$ is a vector of plot level characteristics.

The excluded instrument from the land value equation i.e. equation (3.7) is the log distance to the line. The effect of road is identified by variation in the household's access to road due to its proximity to the north-south line connecting two district headquarters. As discussed above, since the land value equation also includes distance variables and plot characteristics, the instrument is likely to be uncorrelated with the error term ε. Under these conditions, the estimation of β in equation (3.7) using 2SLS, is interpreted as the causal effect of improved access to road on agricultural profits.

3.5 Data

The data for this study come from two major sources. The household information on agriculture and access to infrastructure is obtained from the 2011 Nepal Living Standards Survey (NLSS). It is a nationwide survey collected by the Central Bureau of Statistics with a stratified random sample of 5988 households, out of which 2058 households reside in the rural hill region with farmland. The Survey includes detailed agriculture module with information on all individual plots owned or leased by the household, along with its area, quality, market value, irrigation, and the net rent received by the household if the plot was leased out. Additionally, it collects information on household's access to various facilities, including the nearest paved road, in terms of travel time. Travel time measures are more useful than actual distance calculated using satellite telemetry because of mountainous terrain.

A unique feature of the 2011 NLSS is that it provides GIS information on the location of its households. I use this GIS data to merge the Mountain Environment and Natural Resources' Information System (MENRIS) data for the NLSS sample households. The MENRIS geospatial data is collected by the International Centre for Integrated Mountain Development, ICIMOD, and has information on elevation, foot trails, and river paths for the Hindu Kush Himalaya region including Nepal. I use this to construct all the distance variables including the instrumental variable, and terrain characteristics such as land gradient.

Table 3.1 provides summary statistics for rural households residing in the hills that own a farmland along with their plot characteristics. An average landholdings of a household is only 0.62 ha, with an average plot size of 0.19 ha. 30 percent of the plots are categorized as being suitable for rice cultivation; whereas, almost 70 percent of all plots do not have access to irrigation in any agriculture season.

3.6 Results

The relationship between rent and market value of a plot, as described by the asset-pricing model, can be established by estimating equation (3.4). Table 3.2 presents the estimate of coefficient η_1 in equation (3.4), by using 321 plots that were either rented out in both agricultural seasons or in the wet season. These plots were rented mainly through sharecropping and account for about 5 % of the total plots owned. Their net rents are calculated by summing the rents across both seasons and including the value of in-kind payments, while the costs of inputs provided by the tenants are deducted. The mean rent to value ratio ($\frac{r}{V}$) is 0.055, which can be interpreted as an estimate of the discount rate b.[19]

The OLS estimate of η_1 in column 1 assumes that the log rent is uncorrelated with the error term. Although the estimate is less that one, attenuation bias because of random measurement error in rents could bias the estimate towards zero, as pointed out by *Jacoby* (2000). Therefore, the specifications in columns 2 and 3 instrument rents with plot area and column 3 additionally includes district fixed effects. The estimates, 1.12 and 1.08 respectively, are accurate and not statistically different from unity even at the 10 % level, and therefore the validity of the asset pricing model cannot be rejected.

Table 3.3 estimates of the impact of transportation infrastructure on the economic gains from agriculture using a log-linear specification described by equation (3.7). Because farmland behaves like an asset as proven by the results from table 3.2, higher farm profits because of lower transportation costs should be capitalized in farmland values. Therefore, the total economic benefits of having better road accessibility for agrarian households can be calculated by estimating β, the coefficient of the log travel time to road in equation (3.7), where the dependent variable is the log value

[19] *Jacoby* (2000) also finds the rent to value ratio among rented lands in Nepal using 1996 NLSS to be 0.055.

of farmlands. Only controlling for the size of farmland, the OLS estimate of β in column 1 is statistically significant at the 1 percent level and its value of -0.19 implies that that an increase in the travel time to the nearest road by 10 percent decreases the value of the farmland by 1.9 percent. However, profitable plots may encourage larger road investments or rural roads may be more abundant in relatively more productive regions. The inclusion of plot characteristics in columns 2 and 3, and district fixed effects in column 3, ameliorates the endogeneity problem as long as the location of roads are determined by observed plot characteristics or by unobserved characteristics of the district and not by unobserved characteristics of the specific plots within the district. While the estimate in column 3 implies a strong positive effect of road on farm value, nevertheless, unobserved socio-economic and political factors may influence road allocation within the district, which could bias the above OLS estimates in either direction and necessitate the use of an instrumental variable strategy to overcome the potential endogeneity bias.

Table 3.4 estimates the correlation between the travel time to the nearest road and the distance to the nearest line by using equation (3.5). The sample is restricted to 2058 households living in the rural hills, for which this strategy is most relevant. Panel A presents estimates that use the line constructed based on the algorithm which divides the region into 10 blocks; whereas, the estimates using the line constructed with 12 blocks are shown in Panel B. All the estimates are positive and significant at the 1 percent level, even after controlling for district fixed effects in column 2 and additionally for the distance to the nearest district HQ and the distance to EWH in column 3.

This strong correlation suggests the possibility of using this variable as an instrument in a 2SLS model described by equations (3.6) and (3.7), to estimate the effect of better access to road on land value. The threat to the validity of such a model arises from the potential correlation of the instrumental variable with (1) proximity to district headquarter, (2) closeness to the historical trade corridors, and (3) physical characteristics of the plot. However, NLSS and MENRIS datasets have a wide set of household and plot level information that can be used to construct these variables, which are then included in equations (3.6) and (3.7). The measures of plot characteristics include the distance to river, land gradient, elevation, a dummy indicating whether the plot is suitable for rice cultivation, dummies indicating year round or seasonal irrigation, and dummies indicating types of irrigation. The distance to foot trails picks up unobserved characteristics that are correlated with the proximity to historical trading routes; whereas the distance to the nearest district headquarter and the distance to EWH variables control for the proximity to district headquarter and EWH respectively.

Tables 3.5 and 3.6 estimate the reduced form effect of the distance to the line constructed by using 10 and 12 blocks algorithms, respectively. The specification in column 1 in both tables only includes controls for district fixed effects, farm size, and the proximity to the nearest district headquarter and EWH. According to table 3.5 column 1, an increase in the distance to the line by 10 percent decreases the value of land by 0.34 percent. As discussed above, the distance to the line could be correlated with proximity to the historical trade corridors and physical characteristics of the land, both of which directly affect farmland value. To address this, I control for the shortest distance to the walking trails in columns 2 and 3 and observed plot characteristics in column 3. While many of these control variables have strong explanatory power and sensible signs, the coefficient estimates of the distance to the line do not change considerably in magnitude from previous estimates in column 1 and are statistically significant at the 1 percent level. The results show that plots suitable for rice and plots with year round as well as seasonal irrigation are more valuable. Likewise, an increase in land gradient negatively affects the value of the land; whereas, plots that are closer to the district headquarter or to the foot trails have greater value.

The coefficient estimates on the distance to the line in table 3.5 column 2 and 3 equal -0.039 and -0.042 respectively, both of which are within a standard error of the estimate of -0.034 from column 1. Therefore, it seems unlikely that the instrumental variable– distance to the line– is correlated with unobserved characteristics associated with historical proximity to trade routes and with additional land quality variables that are not captured in the present data set, thereby bolstering the validity of the given instrument.

Tables 3.7 and 3.8 column 3 present IV estimates of road accessibility by using distance to the line as an instrument for the household's travel time to the nearest road. The estimates of the coefficient β from the 2SLS model are -0.234 and -0.263 for the two algorithms and they are statistically significant at the 5 percent level. The results suggest that an increase in the travel time to the nearest road by one percent decreases the farmland value by about 0.25 percent. Table 3.9 shows the results for the first stage of the 2SLS model by estimating coefficient β' in equation (3.6), which suggest that an increase in the distance to the line by one percent raises the household's travel time to the nearest road by 0.17 percent. The F-statistics on the excluded variable from both algorithms in the first stage are large, 24.58 and 20.53, which indicate that the IV estimates are unlikely to be biased due to weak instrument. Moreover, the IV estimates are robust to different parameters of the algorithm used to construct the instrument, suggesting that an improvement in rural road infrastructure has a positive impact on households' economic gains from their agricultural activities.

3.7 Conclusion

Transportation infrastructure plays a crucial role in rural development. It improves agricultural productivity by reducing transaction costs (*Jacoby*, 2000) and promotes efficient labor allocation by stimulating commercialization of agriculture

(*Gollin and Rogerson*, 2010). Since many poor reside in remote locations, it provides a direct means of reaching them, thereby facilitating effective delivery of poverty programs and basic services in health, nutrition, and education. However, evaluating transportation projects poses a considerable challenge because roads are not randomly placed nor it is feasible to randomize by design.

I identify a unique geographic feature of mountains in rural Nepal that allows for a plausibly exogenous variation in the placement of rural roads. The empirical strategy is based on the fact that two important determinants of road construction, administrative need to connect district headquarters and the geographic feature that favors north-south linkages, are likely to be independent of the socio-economic characteristics of households living in the hinterlands. Using the distance to the north-south line that connect two district headquarters as an instrument for road accessibility, I find that reduction in travel time to the nearest road increases farm profits, implying a significant economic gains to rural agricultural households.

Although it establishes a causal link between economic benefits and infrastructure investments, the distributional consequences of such projects are unclear. While remote households have most to gain from improved access, the agricultural profits from farmland are confined to landowners who are usually not among the very poor. Moreover, the mechanisms through which farm profits rise are also not identified. The most likely candidates include a greater use of modern inputs such as chemical fertilizers and improved seeds, an easier dissemination of new farm technologies, and an improved access to urban centers with large demand for agricultural output. These important channels could be identified by employing a similar empirical strategy in the future. Moreover, the benefits from road construction are not only limited to farm profits. *Gollin and Rogerson* (2010) find increase in non-agricultural employment and greater urbanization due to better road network. By promoting factor mobility and improving access to schools, road construction could also raise expected returns to education, prompting households to invest more in human capital. The empirical strategy could easily be extended to evaluate the effect of road on a wider set of economic outcomes of interest including household's education, health, nutrition, and migration decisions, which are important determinants of its welfare and its community's long term economic development.

Figure 3.1: The Physical Map of Nepal

"/Users/sleshshrestha/Dropbox/University of Michigan/Research/Chapter 3-Roads/Figu

Notes: The figure illustrates the ruggedness of the terrain. While the thin southern stretch of the country is flat and uniform, the land surface north of the Terrai plain is characterized by sharp changes in elevation as illustrated by the compactness of contour lines.

Figure 3.2: Historical Trade Routes and Walking Trails

"/Users/sleshshrestha/Dropbox/University of Michigan/Research/Chapter 3-Roads/Figu

Notes: The figure plots the walking trails that span the country. The district headquarters are illustrated by dotted points on the map. The foot trails stretch in a north-south as well as an east-west directions.

Figure 3.3: Lines Constructed by Connecting District Headquarters (Using 10 Blocks Algorithm)

"/Users/sleshshrestha/Dropbox/University of Michigan/Research/Chapter 3-Roads/Figu

Figure 3.4: Five Development Regions of Nepal

"/Users/sleshshrestha/Dropbox/University of Michigan/Research/Chapter 3-Roads/Figu

Figure 3.5: Lines Constructed by Connecting District Headquarters (Using 12 Blocks Algorithm)

"/Users/sleshshrestha/Dropbox/University of Michigan/Research/Chapter 3-Roads/Figu

Figure 3.6: Flow of Rivers and Streams

"/Users/sleshshrestha/Dropbox/University of Michigan/Research/Chapter 3-Roads/Figu

Figure 3.7: Elevation

"/Users/sleshshrestha/Dropbox/University of Michigan/Research/Chapter 3-Roads/Figu

Notes: Blue indicate low elevation land and red indicates high elevation. The elevations changes drastically going from north to south, while the elevation remains fairly constant going from east to west.

Table 3.1: Descriptive Statistics

	Mean	Std. Dev
Panel A: HH Characteristics		
Number of Households	2058	-
Land Area *(Ha)*	0.62	0.62
Number of Plots	3.30	2.19
Household Size	4.76	2.10
Travel Time to Paved Road *(Hrs)*	7.42	13.48
Panel B: Plot Characteristics		
Number of Plots	6785	-
Plot Size *(Ha.)*	0.19	0.25
Plot Value *(Million Rs. per Ha.)*	3.04	17.47
% Suitable for Rice	30.79	46.17
% with Year Round Irrigation	14.62	35.33
% with Seasonal Irrigation	16.34	36.98
% with Tubewell	0.87	9.29
% with Canal	25.78	43.74
% with Pond	0.25	4.99
% with Mixed	0.29	5.42
Elevation *(Meters)*	1215.45	457.63
Land Gradient *(Degrees)*	24.29	11.35
Panel C: Distance Variables *(in Km)*		
Distance to the Line (10 Blocks)	11.51	9.41
Distance to the Line (12 Blocks)	11.88	9.06
Distance to District Headquarters	15.95	8.82
Distance to EWH	51.75	21.18
Distance to River	0.84	0.59
Distance to Foot Trails	1.25	1.26

Notes: The household and plot samples are restricted to rural hill region. The household and plot level data in Panels A and B (except elevation and land gradient) is obtained from 2011 Nepal Living Standard Survey. The rest of the data including distance variables are calculated using MENRIS data.

Table 3.2: Relationship Between Plot Values and Rents

	(1) OLS	(2) IV^{τ}	(3) IV^{τ}
Log Rent (v_1)	0.69**	1.12**	1.08**
	(0.0560)	(0.1122)	(0.1152)
Ho: $v_1 = 1$ *(p-value)*	0.000	0.255	0.439
District Fixed Effect	No	No	Yes
Number of Plots	327	327	327

Notes: This table reports the coefficient of the log of rent from equation (3.4), where the dependent variable is the log of plot market value. If the asset-pricing model is valid and market value of the plot is equal to the discounted future rent stream, then this coefficient $v_1 = 1$. The table also presents the p-value of this test.

τ Log of plot area is the excluded instrument.

* indicates significance at the 5 percent level

** indicates significance at the 1 percent level

Table 3.3: Effect of Road on Land Value (OLS)

	(1)	(2)	(3)
Log Travel Time to Road	-0.190***	-0.173***	-0.146***
	(0.0058)	(0.0056)	(0.0060)
Log Plot Area	0.708***	0.677***	0.670***
	(0.0156)	(0.0150)	(0.0137)
Irrigation:			
Year-Round		0.582***	0.663***
		(0.0978)	(0.0859)
Seasonal		0.088	0.293***
		(0.0876)	(0.0767)
Mode of Irrigation:			
Tubewell		0.824*	0.240
		(0.1788)	(0.1568)
Canal		-0.037	-0.006
		(0.0837)	(0.0730)
Pond		0.452	0.145
		(0.3054)	(0.2640)
Mixed		0.540*	0.401
		(0.2827)	(0.2448)
Plot Quality:			
Suitable for Rice		0.458***	0.211***
		(0.0520)	(0.0462)
District Fixed Effect	No	No	Yes
R Squared	0.301	0.361	0.530
Number of Plots	6750	6750	6750

*significance at the 10 percent level ** significance at the 5 percent level *** significance at the 1 percent level

Table 3.4: Relation Between Travel Time to Road and Distance to the Line

	(1)	(2)	(3)
Panel A: 10 Blocks			
Log Distance to the Line	0.623***	0.603***	0.205***
	(0.0576)	(0.0554)	(0.0682)
Log Distance to District Headquarter			1.174***
			(0.117)
Log Distance to EWH			0.724***
			(0.184)
R Squared	0.055	0.431	0.463
Panel B: 12 Blocks			
Log Distance to the Line	0.592***	0.640***	0.237***
	(0.0601)	(0.0574)	(0.0719)
Log Distance to District Headquarter			1.145***
			(0.1190)
Log Distance to EWH			0.754***
			(0.1840)
R Squared	0.046	0.433	0.463
District Fixed Effect	No	Yes	Yes
Number of HHs	2025	2025	2025

Notes:
* indicates significance at the 5 percent level
** indicates significance at the 1 percent level

Table 3.5: Effect of Road on Land Value Using 10 Blocks (Reduced Form)

	(1)	(2)	(3)
Log Distance to the Line	-0.034*	-0.039**	-0.042**
	(0.0196)	(0.0196)	(0.0187)
Log Plot Area	0.696***	0.696***	0.676***
	(0.0147)	(0.0147)	(0.0141)
Irrigation:			
Year-Round			0.622***
			(0.0874)
Seasonal			0.258***
			(0.0783)
Mode of Irrigation:			
Tubewell			0.287*
			(0.159)
Canal			0.066
			(0.0744)
Pond			0.239
			(0.270)
Mixed			0.552**
			(0.269)
Plot Quality:			
Suitable for Rice			0.200***
			(0.0474)
Land Gradient			-0.169***
			(0.0168)
Log Distance to River			0.007
			(0.0138)
Log Distance to District Headquarter	-0.338***	-0.335***	-0.255***
	(0.0335)	(0.0335)	(0.0327)
Log Distance to Foot Trails		-0.044***	-0.041***
		(0.0116)	(0.0111)
Log Distance to EWH	0.203***	0.191***	0.298***
	(0.0559)	(0.0559)	(0.0601)
Elevation			-0.000***
			(0.0000)
District Fixed Effect	Yes	Yes	Yes
R Squared	0.457	0.458	0.511
Number of Plots	6745	6745	6745

*significance at the 10 percent level ** significance at the 5 percent level *** significance at the 1 percent level

Table 3.6: Effect of Road on Land Value Using 12 Blocks (Reduced Form)

	(1)	(2)	(3)
Log Distance to the Line	-0.0502**	-0.0489**	-0.0449**
	(0.0204)	(0.0204)	(0.0195)
Log Plot Area	0.698***	0.697***	0.677***
	(0.0147)	(0.0147)	(0.0141)
Irrigation:			
Year-Round			0.624***
			(0.0873)
Seasonal			0.263***
			(0.0783)
Mode of Irrigation:			
Tubewell			0.281*
			(0.1590)
Canal			0.065
			(0.0744)
Pond			0.244
			(0.2700)
Mixed			0.558**
			(0.2690)
Plot Quality:			
Suitable for Rice			0.197***
			(0.0474)
Land Gradient			-0.167***
			(0.0168)
Log Distance to River			0.008
			(0.0138)
Log Distance to District Headquarter	-0.322***	-0.325***	-0.251***
	(0.0340)	(0.0340)	(0.0333)
Log Distance to Foot Trails		-0.042***	-0.039***
		(0.0116)	(0.0111)
Log Distance to EWH	0.194***	0.183***	0.295***
	(0.0561)	(0.0561)	(0.0601)
Elevation			-0.000***
			(0.0000)
District Fixed Effect	Yes	Yes	Yes
R Squared	0.457	0.458	0.511
Number of Plots	6745	6745	6745

*significance at the 10 percent level ** significance at the 5 percent level *** significance at the 1 percent level

Table 3.7: Effect of Road on Land Value Using 10 Blocks (IV)

	(1) OLS	(2) Reduced Form	(3) IV$^{\tau}$
Log Travel Time to Road	-0.130***	-	-0.234**
	(0.00627)		(0.1050)
Log Distance to the Line	-	-0.042**	-
		(0.0187)	
Log Plot Area	0.686***	0.676***	0.692***
	(0.0137)	(0.0141)	(0.0152)
Irrigation:			
Year-Round	0.684***	0.622***	0.724***
	(0.0852)	(0.0874)	(0.0959)
Seasonal	0.305***	0.258***	0.334***
	(0.0761)	(0.0783)	(0.0827)
Mode of Irrigation:			
Tubewell	0.212	0.287*	0.169
	(0.1560)	(0.1590)	(0.1640)
Canal	-0.012	0.066	-0.064
	(0.0725)	(0.0744)	(0.0903)
Pond	0.086	0.239	-0.028
	(0.2620)	(0.2700)	(0.2900)
Mixed	0.459*	0.552**	0.380
	(0.2610)	(0.2690)	(0.2770)
Plot Quality:			
Suitable for Rice	0.176***	0.200***	0.159***
	(0.0461)	(0.0474)	(0.0498)
Land Gradient	-0.126***	-0.169***	-0.093**
	(0.0164)	(0.0168)	(0.0374)
Log Distance to River	0.019	0.007	0.028*
	(0.0134)	(0.0138)	(0.0165)
Log Distance to District Headquarter	-0.178***	-0.255***	-0.084
	(0.0264)	(0.0327)	(0.0991)
Log Distance to Foot Trails	-0.032***	-0.041***	-0.024*
	(0.0108)	(0.0111)	(0.0135)
Log Distance to EWH	0.361***	0.298***	0.395***
	(0.0590)	(0.0601)	(0.0689)
Elevation	-0.000	-0.000***	-0.005
	(0.0001)	(0.0000)	(0.0001)
F-Statistics on Instrument			24.58
Probability>F			0.000

τ **Log Distance to the Line** is the excluded instrument. All specifications include district fixed effect.

*significance at the 10 percent level ** significance at the 5 percent level *** significance at the 1 percent level

Table 3.8: Effect of Road on Land Value Using 12 Blocks (IV)

	(1) OLS	(2) Reduced Form	(3) IV$^{\tau}$
Log Travel Time to Road	-0.130***	-	-0.263**
	(0.00627)		(0.1160)
Log Distance to the Line	-	-0.045**	-
		(0.0195)	
Log Plot Area	0.686***	0.677***	0.694***
	(0.0137)	(0.0141)	(0.0156)
Irrigation:			
Year-Round	0.684***	0.624***	0.736***
	(0.0852)	(0.0873)	(0.0989)
Seasonal	0.305***	0.263***	0.342***
	(0.0761)	(0.0783)	(0.0848)
Mode of Irrigation:			
Tubewell	0.212	0.281*	0.157
	(0.1560)	(0.1590)	(0.1670)
Canal	-0.012	0.0649	-0.0779
	(0.0725)	(0.0744)	(0.0944)
Pond	0.086	0.244	-0.0591
	(0.2620)	(0.2700)	(0.2980)
Mixed	0.459*	0.558**	0.358
	(0.2610)	(0.2690)	(0.2830)
Plot Quality:			
Suitable for Rice	0.176***	0.197***	0.154***
	(0.0461)	(0.0474)	(0.0510)
Land Gradient	-0.126***	-0.167***	-0.084**
	(0.0164)	(0.0168)	(0.0408)
Log Distance to River	0.019	0.00842	0.031*
	(0.0134)	(0.0138)	(0.0173)
Log Distance to District Headquarter	-0.178***	-0.251***	-0.058
	(0.0264)	(0.0333)	(0.1090)
Log Distance to Foot Trails	-0.032***	-0.039***	-0.022
	(0.0108)	(0.0111)	(0.0141)
Log Distance to EWH	0.361***	0.295***	0.404***
	(0.0590)	(0.0601)	(0.0713)
Elevation	-0.000	-0.000***	0.000
	(0.0001)	(0.0000)	(0.0000)
F-Statistics on Instrument			20.53
Probability>F			0.000

τ **Log Distance to the Line** is the excluded instrument. All specifications include district fixed effects.

*significance at the 10 percent level ** significance at the 5 percent level *** significance at the 1 percent level

Table 3.9: Travel Time to Road and Distance to the Line (First Stage)

	(1) 10 Blocks	(2) 12 Blocks
Log Distance to the Line	0.176***	0.167***
	(0.0355)	(0.0370)
Log Plot Area	0.061**	0.055**
	(0.0267)	(0.0267)
Irrigation:		
Year-Round	0.432***	0.418**
	(0.1660)	(0.1660)
Seasonal	0.311**	0.291**
	(0.1490)	(0.1480)
Mode of Irrigation:		
Tubewell	-0.483	-0.451
	(0.3040)	(0.3030)
Canal	-0.532***	-0.523***
	(0.1410)	(0.1410)
Pond	-1.129**	-1.142**
	(0.5100)	(0.5100)
Mixed	-0.730	-0.752
	(0.5100)	(0.5100)
Plot Quality:		
Suitable for Rice	-0.162*	-0.153*
	(0.0898)	(0.0899)
Land Gradient	0.323***	0.314***
	(0.0318)	(0.0318)
Log Distance to River	0.090***	0.086***
	(0.0261)	(0.0261)
Log Distance to District HQ	0.730***	0.736***
	(0.0618)	(0.0631)
Log Distance to Foot Trails	0.081***	0.072***
	(0.0211)	(0.0211)
Log Distance to EWH	0.338***	0.347***
	(0.1150)	(0.1150)
Elevation	0.001***	0.001***
	(0.0000)	(0.0000)
District Fixed Effect	Yes	Yes
R Squared	0.445	0.445
Number of Plots	6710	6710

* indicates significance at the 10 percent level ** indicates significance at the 5 percent level
*** indicates significance at the 1 percent level

BIBLIOGRAPHY

BIBLIOGRAPHY

Abadie, A., and J. Gardeazabal (2003), The economic costs of conflict: A case study for the basque country, *The American Economic Review*, *93*(1), 113–132.

Angelucci, M., G. DeGiorgi, M. A. Rangel, and I. Rasul (2009), Family networks and school enrollment: Evidence from a randomized social experiment, *NBER Working Paper*.

Angrist, J., and V. Lavy (1999), Using maimonides' rule to estimate the effect of class size on scholastic achievement, *Quarterly Journal of Economics*, *114*(2), 533–575.

Angrist, J. D., and G. W. Imbens (1995), Two-stage least squares estimation of average causal effects in models with variable treatment intensity, *Journal of the American Statistical Association*, *90*(430), 431–442.

Arnott, R. J., and J. E. Stiglitz (1981), Aggregate land rents and aggregate transport costs, *The Economic Journal*, *91*(362), 331–347.

Atack, J., F. Bateman, M. Haines, and R. Margo (2009), Did railroads induce or follow economic growth? urbanization and population growth in the american midwest, 1850-60, *NBER Working Paper No. 14640*.

Banerjee, A., E. Duflo, and N. Qian (2009), On the road: Access to transportation infrastructure and economic growth in china.

Banskota, P. (1994), *The Gurkha Connection: A History of the Gurkha Recruitment in the British Indian Army*, Nirala Publications.

Becker, G. (1965), A theory of the allocation of time, *Economic Journal*, *75*, 493–517.

Becker, G. (1981), *A Treaties on the Family*, Harvard University Press.

Behrman, J. R., R. Pollak, and P. Taubman (1982), Parental preferences and provision for progeny, *The Journal of Political Economy*, *90*(1), 52–73.

Beine, M., F. Docquier, and H. Rapoport (2006), Brain drain and human capital formation in developing countries : Winners and losers, university catholique de Louvain, Departement des Sciences Economiques.

Bertrand, M., E. Duflo, and S. Mullainathan (2004), How much should we trust differences-in-differences estimates?, *Quarterly Journal of Economics*, *119*(1), 249–275.

Bhagwati, J., and K. Hamada (1974), The brain drain, international integration of makets for professionals and unemployment, *Journal of Development Economics*, *1*, 19–42.

Bommier, A., and S. Lambert (2000), Education demand and age at school enrollment in tanzania, *The Journal of Human Resources*, *35*(1), 177–203.

Burgess, R., and R. Jedwab (2010), Our turn to eat: The political economy of roads in kenya.

Caplan, L. (1995), *Warrior Gentlemen: Gurkhas in the Western Imagination*, Berghahn Books.

Chalamwong, Y., and G. Feder (1988), The impact of landownership security: Theory and evidence from thailand, *World Bank Economic Review*, *2*, 187–204.

Chand, S., and M. A. Clemens (2008), Skilled emigration and skill creation: A quasi-experiment.

Docquier, F., O. Lohest, and A. Marfouk (2007), Brain drain in developing countries, *World Bank Economic Review*, *21*(2), 193–218.

Donaldson, D. (2010), Railroads and the raj: The economic impact of transportation infrastructure.

Duflo, E., and R. Hanna (2005), Monitoring works: Getting teachers to come to school, *NBER Working Paper*, *11880*.

Fafchamps, M., and R. V. Hill (2005), Selling at the farm-gate or travelling to market, *American Journal of Agricultural Economics*, *87*(3), 717–734.

Foster, A. D., and M. R. Rosenzweig (1996), Technical change and human-capital returns and investments: Evidence from the green revolution, *The American Economic Review*, *86*(4), 931–953.

Gollin, D., and R. Rogerson (2010), Agriculture, roads, and economic development in uganda, *NBER Working Paper No. 15863*.

Gronau, R. (1977), Leisure, home production and work: The theory of the allocation of time revisted, *Journal of Political Economy*, *85*(6), 1099–1123.

Haque, N. U., and S.-J. Kim (1995), Human capital flight: Impact of migration on income and growth, *IMF Staff Papers*, *42 (3)*, 577–607.

Hitchcock, J. T. (1966), *The Magars of Banyan Hill*, Holt, Rinehart, and Winston.

Jacoby, H. G. (2000), Access to markets and the benefits of rural roads, *The Economic Journal*, *110*(465), 713–737.

Jacoby, H. G., and B. Minten (2008), On measuring the benefits of lower transport costs, *World Bank Policy Research Working Paper 4484*.

Jensen, R. (2010), The (perceived) returns to education and the demand for schooling, *The Quarterly Journal of Economics*, *125*(2), 515–548.

Kochar, A. (2004), Urban influences on rural schooling in india, *Journal of Development Economics*, *74*, 113–136.

Kremer, M., E. Miguel, and R. Thornton (2009), Incentives to learn, *Review of Economics and Statistics*, *91*(3), 437–456.

Lee, J. (2008), Sibling size and investment in children's education: An asian instrument, *Journal of Population Economics*, *21*(4), 855–875.

Manski, C. F. (1993), Identification of endogenous social effects: The reflection problem, *The Review of Economic Studies*, *60*(3), 531–542.

Mansuri, G., S. A. Shrestha, H. Winkler, and M. Yanez-Pagans (2012a), A plot of my own: Property rights and economic mobility in rural uganda.

Mansuri, G., S. A. Shrestha, H. Winkler, and M. Yanez-Pagans (2012b), Health or death? health shock of an income earner and economic mobility in rural pakistan.

McKenzie, D., and H. Rapoport (2006), Can migration reduce educational attainment? evidence from mexico, *World Bank Policy Research Working Paper 3952*.

Meyer, W. (2008), Nepal-roads to prosperity partnership results, *Tech. rep.*, SDC Swiss Agency for Development and Cooperation.

Michaels, G. (2008), The effect of trade on the demand for skill-evidence from the interstate highway system, *Review of Economics and Statistics*, *90*(4).

Miyagiwa, K. (1991), Scale economies in education and the brain drain problem, *International Economic Review*, *32*(3), 743–759.

Mountford, A. (1997), Can a brain drain be good for growth in the source country?, *Journal of Development Economics*, *53(2)*, 287–303.

Newman, J. L., and P. J. Gertler (1994), Family productivity, labor supply, and welfare in a low income country, *The Journal of Human Resources*, *29*(4), 989–1026.

Oyelere, R. U. (2009), Understanding low average returns to education in africa: The role of heterogeneity across education levels and the importance of political and economic reforms.

Pitt, M. M., M. R. Rosenzweig, and M. N. Hassan (1990), Productivity, health, and inequality in the intrahosuehold distribution of food in low-income countries, *The American Economic Review*, *80*(5), 1139–1156.

Qian, N. (2009), Quantity-quality and the one child policy: The only-child disadvantage in school enrollment in rural china, *NBER Working Paper*.

Rathaur, K. R. S. (2001), British gurkha recruitment: A historical perspective, *Voice of History*, *16(2)*, 19–24.

Rosenzweig, M. R., and K. I. Wolpin (1980), Testing the quanity-quality fertility model: The use of twins as a natural experiment, *Econometrica*, *48*(1), 227–240.

Schultz, P. T. (2004a), School subsidies for the poor: Evaluating the mexican progresa poverty program, *Journal of Development Economics*, *74*(1), 199–250.

Schultz, T. P. (2004b), School subsidies for the poor: Evaluating the mexican progresa poverty program, *Journal of Development Economics*, *74*(1), 199–250.

Schultz, T. W. (1975), The value of the ability to deal with disequilibria, *Journal of Economic Literature*, *13*(3), 827–846.

Shrestha, B., and S. Jain (1977), *Regional Development in Nepal*, Development, New Delhi.

Shrestha, N. R. (2001a), *The Political Economy of Land, Landlessness and Migration in Nepal*, Nirala Series.

Shrestha, N. R. (2001b), *The Political Economy of Land, Landlessness and Migration in Nepal*, Nirala Series.

Shrestha, S. A. (2011), Effects of educational returns abroad on domestic schooling: A british gurkha amry experiment.

Shrestha, S. H. (2004), *Economic Goegraphic of Nepal*, Educational Publishing House.

Stark, O., C. Helmenstein, and A. Prskawetz (1997), A brain gain with a brain drain, *Economic Letters*, *55*(2), 227–234.

Stiller, L. (1976), *The Silent Cry*, Sahayogi Prakashan, Kathmandu.

Strauss, J., and D. Thomas (1995), Human resources: Empirical modeling of household and family decision, *Handbook of Development Economics*, *3*, 1885–2005.

Upadhyaya, S. K. (2010), Upland poverty in nepal: The role of environment, *Tech. rep.*, Institute for Integrated Development Studies.