The Economics of Inequality and Human Capital Development: Evidence from Nepal

Ram Prasad Mainali

March, 2014

A thesis submitted to

the Academic Faculty

by

Ram Prasad Mainali

In partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Economics

City University London

London, United Kingdom

March, 2014

Contents

List of Figures

List of Tables

Acknowledgement

I would like to express my gratitude to the Department of Economics, City University, London and its staffs for providing me financial, academic and technical supports particularly for the award of studentship that provided financial support for this study. Special thanks go to the coordinator of PhD program for selecting me in this program. I am also indebted to the government of Nepal for granting me a study leave. Personal supports by all my seniors and colleagues in the Ministry of Finance (MoF) undoubtedly deserve my appreciation.

I am grateful to my principle supervisor, Professor Gabriel Motes-Rojas and second supervisor, Professor Saqib Jafarey, for their thought provoking intellectual guidance and in-depth discussion in absence of which this study would not have been completed. Excellent advices, support and friendship of my third supervisor, Dr Victoria Serra-Sastre, has been invaluable on both an academic and a personal level, for which I am extremely grateful.

My gratitude goes to all academic staffs and fellow students in MPhil and PhD programs at the City University. I would also like to take this opportunity to express my sincere thanks to all academic participants of departmental seminars, City University; PhD Res conference, Queen Mary University; Eurasia Business and Economic Society (ESBS) conference, Istanbul; European Economic Association (EEA) conference, Malaga and British Nepal Academic Council (BNAC) conference, university of Edinburg for their valuable comments on my research.

I may offer my sincere thanks to my wife Prerana Mainali for not only taking all the household responsibilities and releasing me from household chores but also for her morale support. I am certainly obliged to my daughter Pratyuja Mainali and son Raunak Mainali for taking their own care and continuing study in absence of my proper guidance, which was certainly necessary for them particularly at this young stage. My brothers and sisters have given me their unequivocal support throughout my study for which I am extremely grateful.

I am equally indebted to all the informants who provided precious insight over the years of study.

Any errors or inadequacy that may remain in this work is entirely my own responsibility.

Declaration

- This dissertation was written by Ram Prasad Mainali based on work undertaken by him in collaboration with professor Gabriel Montes-Rojas, professor Saqib Jafarey and Dr Victoria Serra-Sastre at Economics Department, School of Social Science, City University London.
- The second and third chapters were made in collaboration with professor Gabriel Montes-Rojas and professor Saqib Jafarey. The fourth chapter was made in collaboration with professor Gabriel Montes-Rojas and Dr Victoria Serra-Sastre.
- This work has not been accepted for any previous degree.
- This work has not been published.

Abbreviation

AIP: Ability to Pay.

BNAC: British Nepal Academic Council.

CBS/N: Central Bureau of Statistics/Nepal.

CHIS: California Health Interview Survey.

CRS: Constant Return to Scale.

CSLCG: Comprehensive Survey of Living Condition in Japan.

DfID: Department for International Development.

EEA: European Economic Association.

EHCS: Essential Health Care Service.

ESBS: Eurasia Business and Econometric Society

FOC: First Order Condition.

GC: General Caste.

GDP: Gross Domestic Product.

GoN: Government of Nepal.

GOPM: Generalised Ordered Probit Model.

GPI: Gender Parity Index.

HDI: Human Development Indicator.

HIS: Health Interview Survey.

IFPRI: International Food Policy Research Institute.

ILO: International Labour Organization.

INGO: International Non-Governmental Organization.

IV: Instrumental Variable.

LDCs: Least Developed Countries.

LSMSS: Living Standard Measurement Survey Strategy.

LTHP: Long Term Health Policy.

MDGs: Millennium Development Goals.

MENA: Middle East and North Africa.

MoF: Ministry of Finance.

MoHP: Ministry of Health and Population.

MP : Marginal Productivity.

NCO: National Classification of Occupation.

NCPM: Nepalese Communist Party Maoist.

NFHS: National Family Health Survey.

NGO: Non-Governmental Organization.

NHSP-IP: Nepal Health Sector Program-Implementation Plan.

NLSS: National Living Standard Survey.

NLTHP: Nepal's Long-term Health Plan.

NPC/N: National Planning Commission/Nepal.

NSDUH: National Survey on Drug Use and Health.

OECD: Organization for European Cooperation for Development.

OLS: Ordinary Least Square.

OOP: Out-of-Pocket.

OPM: Ordered Probit Model.

PHCH: Primary Health Care Service.

PRSP: Poverty Reduction Strategy Paper.

PSU: Primary Sampling Unit.

RHIS: Rural Health Insurance Scheme.

SAARC: South Asian Association for Regional Cooperation.

SAH: Self Assessed Health.

SC: Scheduled Caste.

SES: Socio Economic Status.

SIC: Standard Industrial Classification.

UIIS: Urban Inpatient Insurance Scheme.

UNDP: United Nation Development Program.

UNESCO: United Nations Educational, Scientific and Cultural Organization.

UNICEF: United Nations Children's Fund.

USA: United State of America.

WB: World Bank.

WHO: World Health Organization.

Abstract

This thesis has three pieces of empirical studies that analyse economic inequality across social groups (castes and gender) and its impact on human capital endowments in developing countries with particular reference to Nepal. Three aspects of inequalities have been examined: disincentive in educational attainment in female arising from labour market discrimination, disproportional representation of low-caste workers in better jobs and inequity in health care utilisation and health outcomes across castes. This study contributes to the literature of economics by developing a new theory and extending existing econometric models in analysing economic inequality across social groups.

The first piece of research examines the impact of marital anticipation on female education in the presence of labour market discrimination. It develops a theoretical model for jointly determining the age at marriage and female education. The model hypothesizes that as females are not rewarded in the labour market as much as men are; married women are encouraged to engage in household work as a result of the intra-household division of labour in their marital union. Thus, parental anticipation of this effect affects their daughter's age at marriage and can influence investment in girls' schooling. It then estimates the causal effect of age at marriage on education in light of the theoretical model using household data from Nepal. In order to control for potential reverse causality this study uses variation in cultural norms regarding dowry and differences in average age of female marriage among ethnicities and regions as instrumental variables. The econometric results confirm that the gender gap in education is significantly affected by cultural practices that favour early marriage and that increasing girls' marriage age by one year would produce on average .4 year increment in women's schooling.

The second study examines the sources of wage differentials across castes in Nepal by employing an extended form of Oaxaca decomposition methodology. This study shows that, in countries such as Nepal which have imperfect labour market for both goods and services, the conventional Oaxaca decomposition methodology fails to estimate the source of wage differential precisely. Thus, it estimates the sources of caste wage differential by using an extended

model of occupational choice, firm size distribution and the interaction between these two along with the conventionally used measures of human capital endowments. Furthermore, it examines the caste differences in the likelihood of access to jobs in large firms. Results indicate that the lack of access to better paying occupations and larger firms have significant impact on caste wage differential. In addition, it evaluates the impact of government policy interventions on caste wage differential in Nepal and shows that the government policy of 'affirmative action' has not yet been effective in narrowing down the caste wage differential in the labour market.

The third piece of research evaluates caste- inequity in health care utilisation and examines the determinants of Self-Assessed Health (SAH) status across castes. It argues that societies with a caste based social stratification in the past perpetuate health sector inequity via inferior social capital in historically discriminated-against castes. Additionally, this study evaluates the effectiveness of the health policy of the Government of Nepal (GoN) that aims to support poor and vulnerable people and consequently to promote equity. The empirical evidences reveal that in a social setting of caste classification the historically discriminated-against caste groups, low castes, face both inequities in health care utilisation and health outcomes. Nevertheless, both types of inequities are decreasing over time this study did not find explicit evidences in favour of the effectiveness of government health policy intervention.

Chapter 1

Introduction

1.1 Introduction

Economic inequality can broadly be defined as the unequal distribution of economic benefits among equally prospective economic agents. This is widely studied in the field of labor economics while examining labour market discrimination across social groups.

Discrimination in labour market is defined as paying differently for workers from different groups, i.e., race, gender, caste, ethnicities etc. with identical productivity characteristics. These concepts emerge from the theories of taste discrimination, whereby, employers directly hold preferences about group background of their employees (Becker 1957, 1971) and statistical discrimination, whereby employers have incomplete information about workers' productivity and statistical priors about how productivity varies with social groups (Arrow, 1971; Phelps, 1972; Akerlof, 1984). Thus the literatures in the labour market discrimination are grounded on either one of these theoretical frameworks. However, both of these theories are based on neoclassical model that implies that the competition will lead to the elimination of discrimination in the long-run. In contrast, empirical evidence show that the

wage differentials across social groups continues even in the long-run indicating the weakness of both theories in explaining the persistence nature of labour market discrimination that perpetuates income inequality in societies.

Being motivated by these limitations in the literature of labor economics this research argues that discrimination persists via *pre-market* effects accede to the deficiency in human capital endowment in discriminated-against social groups that in turn reinforces income inequality among them. Additionally, labor market discrimination is not only a form of economic inequality that different groups in a society face in their practical lives. Other sources of discriminations, for instance, unequal distribution of health care services can produce inequality in human capital endowment and underpin income inequality indirectly. Therefore, the perseverance of income inequality should be studied in the context of deficiency in human capital in disadvantaged social groups.

The conceptual framework depicted in the figure 1.1 below shows that deficiency in human capital stock in discriminated-against groups can be produced via two channels. First; the disadvantaged groups might have fewer incentives to invest in human capital because of its low return. Secondly, a relatively low level of income, which can partly be attributed to the current market discrimination, undermines their ability to invest in human capital, i.e., to invest in health or in education. In other words, current market discrimination reinforces negative effects on ability to invest as well as the willingness to invest in human capital endowment in discriminated-against groups. It implies that the labor market discrimination can lead to a vicious circle of income inequality via *pre-market* effect even if there is no longer practice of current market discrimination by employers.

Nonetheless, there is a substantial lack in the empirical literature that attempt to analyse such linkages. That is how discriminated-against social groups end up in

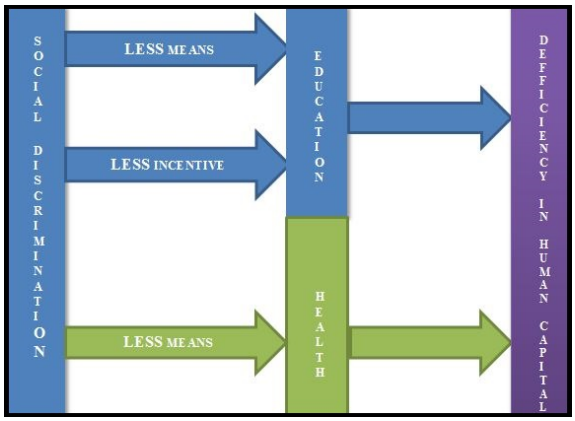

Figure 1.1: Conceptual framework

an inferior human capital endowment that impede them in achieving "fair share" of socioeconomic outcomes at present as the effect of past discrimination. The "fair share" can be viewed in terms of proportionate representation in white collar jobs, better occupation , access to larger firms or obtaining better health status etc. which ultimately affects their economic wellbeing via the low level of socioeconomic status. This study aims to examine these issues. This section, thus deals with the rationale for study, its objectives and samples etc.

1.2 Rationale for study

Over the last few decades, the notion of empowering socially disadvantaged groups has dominated the public policy agenda particularly in developing countries. Social planners in these countries have paid special attention in narrowing down the gender gaps in education and income inequality based on social identity such as caste, race and ethnicities while designing their development policies. These issues had been addressed not only in country-specific context but also at the international arena. For instance, the United Nation Development Program (UNDP), in its Millennium Development Goals (MDGs), vowed to attain gender parity in education by

2015. All member countries have applied their own action plans to achieve MDGs. Some countries have even provided cash incentive for girl's enrolment in school (for instance: Bangladesh). Despite these, gender gap in education seems to remain significant even after 2015.[1]

Additionally, international development partners such as the World Bank (WB) and Department for International Development (DfID) explored caste/ethnic and regional disparities in socioeconomic status (SES) in developing countries (Bennet, 2005) and showed their concern to empower historically disadvantaged social groups. Acknowledging this, developing countries persuaded the policy of affirmative action to enrich disadvantaged groups' SES. The affirmative actions include the assurance of proportionate representation in different levels of political structure, target-group allocation of national budget along with the provision of *qota* in public sector jobs for disadvantaged groups. However, analogous to the persistence gender gaps in education the inequality in SES across castes remains significant in historically caste-based societies.

These indicate that there might be additional factors that marginalized the expected outcomes of such policy interventions. Therefore, the main purpose of this study is to explore additional underlying factors that perpetuate gender gaps in education and caste inequality in income with particular reference to labour market outcome. In particular, marital prospect has been analysed in explaining persistence of gender gap in education whereas the group difference in access to better paid jobs, proxied by larger form, has been examined as the casual factor of caste wage differential.

Finally, it examines inequity in health care utilisation and evaluates the determinants of health status across castes in Nepal. A special attention is given to

[1]UNESCO predicts 17.3% women in South Asia 23% in sub-Saharan Africa will remain illiterate by 2020 while same measures stand at 8.1% and 11.3% respectively, for male (UNESCO 1993).

see whether low-castes end up in inferior health outcomes relative to the dominant caste. Hence, this research aims to find answers for the following research questions.

- Does the marital anticipation affect female educational attainment ?
- What is the role of difference in access to better paid jobs, larger firm, in producing caste differential in labour market outcome ?
- Does inequity in health care utilisation across castes prevail in historically caste-based society ?
- Do historically discriminated-against castes face inferior health outcomes relative to the dominant caste ?

1.3 Research objectives

This research has following objectives

- To extend the literature by developing a new theoretical framework to explain the impact of marital anticipation on female education.
- To offer an empirical test for newly developed theoretical framework.
- To explore additional factors of caste wage differential by extending existing methodologies.
- To examine the inequity in health care utilisation across castes.
- To examine the impact of caste on health outcomes.

1.4 Data

This research employs nationally representative household survey data from Nepal for various periods. Details of the samples are described in respective chapters.

1.5 Research scheme

The research is divided into following three chapters.

Chapter One : The Impact of Marital Anticipation on Female Education: Theory and Evidence from Nepal.

Chapter Two: Earnings and Caste: An Evaluation of Caste Wage Differentials in the Nepalese Labour Market.

Chapter Three: Caste Inequity in Health Care utilisation and The Impact of Caste on Health Outcomes: Evidence from Nepal.

Chapter 2

The Impact of Marital Anticipation on Female Education: Theory and Evidence from Nepal (joint with professor Gabriel Montes-Rojas and professor Saqib Jafarey)

2.1 Introduction

Education is not only a human right but also an input into economic development. In particular female education has been found to be positively associated with many desirable socioeconomic outcomes such as lower fertility rates, healthier and better educated children and greater female labour force participation (Bayisenge,2010).

Despite this, females fall behind males in educational attainment. The problem is particularly severe in South Asia, where the Gender Parity Index (GPI), which measures the female to male ratio in education, is significantly low. In this region, the GPI for pre-primary, primary, secondary and upper secondary level enrolment is, respectively, .98, .86, .83 and .75 whereas the global average for the same measures stands at .99, .93 , .93 and .92, respectively. All countries in this region, apart from Sri Lanka and Maldives, lag far behind from the global average of girl's school enrolment ratio and this is increasing on level of schooling (see Table 2.3).

Existing empirical studies on the gender gap in education can be divided into two strands. The first focuses on household characteristics, such as economic status and parental education, and how these influence gender preferences in schooling. In general these studies find that poverty, lack of social security, credit markets and low levels of parental education all contribute to gender biases in educating children (see Cameron and Worswick, 2001; Sawada, 1997). The second strand of empirical studies takes into account gender differences in labour market outcomes. Using data from Uttar Pradesh, India, Kingdom (1998) found that girls face lower economic rate of returns to education. This is not a robust finding, however. For example, Aaslam (2009) decomposes the gender-specific returns to education in the Pakistani labor market between a labour market effect which captures discrimination on the part of employers and a pure education effect. He found that while the pure returns to education were significantly higher for females, the overall market returns were higher for males. Other authors found returns to education to be higher for females in Indonesia (Behrman and Deolalikar (1995)) and Bangladesh (Asadullah (2006)). Moreover, Munshi and Rosenzweig (2006) argue that when a traditional institution such as the Indian caste system interacts with modern labour market institutions, the results can favour female education among the lower castes, since

males belonging to these castes are more likely to face occupational constraints imposed by the caste system. Using panel data from Bombay over the period 1982-2002, they found that girls were more likely to be enrolled in English-medium schools that prepare students for working in white-collar jobs, while boys were more likely to be enrolled in schools where a local language was used as the medium of instruction and which ultimately led to their students following a traditional occupational path. Thus, it is not clear whether low female education in South Asia can be completely explained by expected lower returns to female education.

An important unexplored dimension of the gender gap in education is marriage. In Asian, particularly South Asian, cultures parents consider a daughter's marriage to be one of the family's main milestones and start planning for it years in advance.[1] Not much attention has been paid to how parental plans regarding a daughter's marriage might influence their decisions regarding her education. On the theoretical side, some papers have argued that the prospect of marriage alone biases parents against educating their daughters. Lahiri and Self (2007) analyse the impact of patrilocality in post-marital living arrangements on female education. Patrilocality, which is especially widespread in South Asian countries, leads to the anticipation that a daughter's future earnings will accrue to her in-laws' household rather than her natal household and this discourages investment in her education. Jafarey (2011) argues that due to gender wage inequality in labour markets, the marital division of labor will encourage lead to female's shouldering a larger share of responsibility for housework and the anticipation of this effect will discontinuously lower the value of her education relative to her hypothetical single self.

Following Becker's seminal work (Becker, 1973) on the theory of marriage, there has developed an empirical literature that treats marriage itself as an endogenous

[1] A common metaphor in Urdu, the main language of Pakistan, for someone being sound asleep is "he/she is sleeping like he/she has just married off all his/her daughters".

variable and studies its determinants. However, to my knowledge, only three empirical papers, Mensch et al. (2005), Brien and Lillard (1994) and Field and Ambrus (2008), have so far studied the interaction between marriage and female education. The first study evaluates the effect of expansion in schooling and urbanization on the proportional increase in age at first marriage. However, this study does not consider the effect of age at marriage on education. The second study develops a theoretical framework for joint estimation of education, marriage and first conception. It shows how educational attainments and marital behaviour can influence each other, i.e, the impact of education on the probability of getting married and the probability of enrolment in upper grades after marriage.

Field and Ambrus (2008) look at the effect of early marriage on female schooling and other adult outcomes in Bangladesh. This paper argues that in impoverished and culturally traditional societies parents have an incentive to marry their daughters young as a form of protection against economic vulnerability. The age of menarche imposes a constraint on how early girls can be married. Therefore, the authors use the timing of menarche as an instrument in identifying the impact of early marriage on female schooling. They find that early marriage significantly lowers female schooling and that each year's delay in marriage would increase female schooling by 0.22 years.

Although the study carried out by Field and Ambrus (2008) is an important step in isolating the effects of age at marriage on female education, the relevance of menarche as an instrumental variable for age at marriage is limited to social settings in which child marriage is prevalent. While this might be true of Bangladesh, it is not necessarily true even of other South Asian countries.[2] Moreover, both of

[2]A report published by the United Nations Children's Fund (UNICEF) shows that Bangladesh has a considerably higher ratio of females marrying below the age of 15 years than the other countries in this region. For example, in Bangladesh, approximately 30% of married females from the age group 20-24 are married bellow the age of 15 years while in India the corresponding figure

these studies (Brien and Lillard, 1994 and Field and Ambrus, 2008) look into the *direct* effect of marriage on female education. The concern of this study is with the *indirect* disincentive to female education that marriage can exerts via its implied division of labour. It can be observed that once a woman gets married her burden of household work increases. Therefore, even a female married after the normal age for a particular levels of schooling, the sooner she plans to marry after reaching that age, the less likely that she will attain that level in the first place.

This study contributes to the analysis of the above effect by first outlining a theoretical framework for jointly determining female education and planned age at marriage. The framework is based on Jafarey (2011), in which gender wage inequality is shown to lead to both a direct discount on female education and an indirect one following from the marital division of labour which allocates women to spend relatively more time in housework and men in market work. It is shown that the indirect discount decreases with the anticipated age of marriage of a female. In addition, the age at marriage can itself depend on individual and cultural factors, such as a female's ability to benefit from schooling and/or cultural expectations regarding an ideal age for her to marry.

Second, this study tests the causal effect of age at marriage on female education using data from a household survey in Nepal. Since the theoretical framework suggests that females may select into early marriage on the basis of idiosyncratic and unobservable differences in ability, it cautions that least-squares estimates will be potentially biased. Therefore an instrumental variables strategy is used.

Nepal is well suited for this study because it has considerable variation in age at first marriage across ethnic groups and communities. In particular, members of the Maithili community, which is concentrated in the regions bordering India, have

is 18%. The same measure stands at less than 10% in Nepal, Pakistan and Sri-lanka (UNICEF, 2011).

been identified by ethnographers as practicing an extreme form of dowry culture which leads them to marry their daughters young. Table 2.5 shows a comparison of dowry cost as well as wedding expenses in current year between Maithili and other communities in Nepal. This shows that average dowry paid by Maithili households is almost three times larger than that of the dowry paid by non-Maithili households. The ratio of dowry payment and wedding expenses to household wealth, captured by the price of land holding and household income are also significantly high within the Maithili community. For instance, rural Maithili households pay dowry approximately 25% of their price of land holding whereas the same payment stands at less than 3% for rural non-Maithili households. Similar trends can be observed in the ratio of dowry cost and wedding expenses with household income. This further implies that this particular community fare higher burden of weeding cost relative to their wealth status than the rest of communities in Nepal. There is also a significant rural/ urban variation in dowry cost within Maithili communities.

The particular dowry practice, locally known as *Tilak Pratha*, which Maithili communities practice in an extreme form, is effectively a groom price. In order to find a suitable groom, a Maithili girl requires not just physical attributes and a suitable family background, she also needs to pay *Tilak* money, which increases along with the educational qualification and social standing of the boy (Das, 2009). One reason for their strong adherence to this practice is the geographical and cultural proximity of Maithilis to India, where dowries and groom prices are more prevalent than in other communities of Nepal. Empirical findings from India also suggest a positive correlation between the size of the dowry and the socioeconomic standing of the prospective husband (Halli, 2003). The result of *Tilak Pratha* is that parents try to get their daughters married as soon as possible because older girls are more likely to match with more mature and well-educated boys, putting upward pressure

on the amount of dowry. The survey data used in this analysis shows that 65% of Maithili girls were married by the age of 16, compared with 42% of non-Maithili girls (see Table 2.8). These differences are significant even at the 1% level, and are *prime facie* evidence that *Tilak Pratha* influences marital behaviour in Nepal within the Maithili community.

One of instrumental variables will therefore be a dummy variable indicating membership of the Maithili community.[3] However, since there is the possibility that the Maithili community's cultural attitudes are biased against female education (relative to those of other communities), this study uses a second instrument, namely the average age at marriage within the respondent's ethno-regional group. The second instrument captures joint ethnic and regional variation with regards to marital behaviour and is less likely to be affected by the prevailing cultural attitudes of the respondent's own reference group towards female education. Result shows that reducing age at marriage by one year reduces female education by approximately 0.4 years of schooling.

The rest of the study is organized as follows. Section 2.2 reviews previous literature related to the female education. Section 2.3 describes the theoretical framework followed by data and descriptive statistics in Section 2.4. Section 2.5 explains the econometric model. Section 2.6 reports the empirical results. The last section concludes.

[3]Dowry payment itself can not be used as an instrumental variable since survey provides information regarding dowry only for the current year. Similarly, a comparison of results by using rural and urban Maithilis as instruments, as there might be significant difference in the adherence of dowry culture between rural and urban Maithilies, is not possible due to a negligible proportion of them living in urban areas. Survey data used in this study shows that 97% of them reside in rural areas.

2.2 Literature review

2.2.1 Investment in human capital endowment

Human capital refers to skills or knowledge acquired by a person that enhances his or her economic value. Shultz (1961) has pointed out five different categories of human capital investment: health facilities and services, on-the-job training, education (Formal), study programs for adults and migration. However, since this study focuses on female education discussion below concentrates on the investment in formal education.

The basic theory of human capital production pioneered by Becker (1962) explains individual choice of investing in human capital endowment analogously to the behaviour of profit maximizing forms. In other words, an individual continues to acquire extra units of human capital until the marginal cost of obtaining it equalizes the marginal benefit that it produces. Costs of education include direct cost such as tuition fee, books stationaries, etc. and indirect cost which is the opportunity cost for not allocating time in the labour market.

Since the cost of acquiring education is an upfront cost a discounted present value of life time earning is compared when making an investment decision under this theoretical framework. Thus, this model shows that an individual's preference of acquiring an extra unit of human capital declines with the discount rate whereas it rises with the ability parameter and expected time of its utilisation over life time. This model also shows an increase in lifespan of an activity would positively affect the rate of return. Therefore, a young person is likely to invest more in education relative to the old person because the former can collect the returns longer. This model explains investment decisions on different points of time but do not show how the accumulation of human capital varies over the life cycle. Additionally, it does

not incorporate depreciation in human capital.

An extended model, the production of human capital and the life cycle of earnings (Ben-Porath, 1967) shows the time path of human capital accumulation incorporating a depreciation component into the model. It shows that not only the human capital but also the investment declines with age. An increase in interest rate reduces the production of human capital and investment along with the rate which investment declines. However, if rental price of human capital and unit price of other inputs rises in the same ratio it leaves quantity of human capital produced unaffected. On the other hand, production of human capital will rise if rental price rises more than the cost of other inputs and *vice versa*.

2.2.2 Human capital theory and female education

As stated in the previous paragraph, a standard human capital model exhibits that individuals decision to attain extra units of human capital (education) depends on the present value of future earnings. That is the added unit of human capital complements to the future earning spectrum of an individual. However, this model does not explicitly incorporate how earnings itself can be influenced across individuals even with similar levels of educational endowment. For instance, Becker (1975) claimed that educational benefits to women is determined by family earnings rather than her own earning due to a significantly less labour force participation by married women (Yokozeki, 2000). Additionally, Chiswick (1970) argues that female labour force participation is not independent as of men and single women but is strongly influenced by other factors such as husband's income, number of children and their age distribution. Therefore, this author excludes married women and students in an empirical analysis of returns to education.

Another factor that undermines female education relative to men is the gender

difference in access to resources. Lerner et al. (1997), while analysing gender gaps in entrepreneurial performances, indicate a lack of pre-requisites (human capital as well as financial capital) in women to achieve success because of their fewer access to resources compared to men. Female can face the lack of resources because of their expected career path as well as the rate of labour force participation. Both of these variables can largely be influenced by the socio-specific cultural context. Therefore, gender gaps in human capital vary across regions depending in the cultural contexts (Rosenbusch et al., 1994). In a society where women are expected to work more in domestic chores are likely to have less human capital endowment than where they work more in the labour market. Woodhall (1973) describes females' education as either a form of consumption or unprofitable form of investment.

This analysis leads to a conclusion that female's education can not merely be explained within the framework of the traditional standard theory of human capital production. In other words, the concept of human capital investment may not apply to women as it does to men. Two compelling approaches, intrahousehold allocation of resources and gender difference in returns to education are generally referred in deriving economic explanation of gender gaps in education. Household income influences parental decision to invest in children's education especially when families face credit constraints (Glewwe and Jacoby, 1994).

2.2.3 Parental preferences

Various types of parental characteristics and its linkage with their preference in allocating household resources in children's education have been analysed in literature to explain gender gaps in education. Altruistic parents decide whether to invest in boys' or in girls' education taking into account the innate ability associated with their respective children. If the cost of educating children with different ability is the

same parents invest more on less able children in order to compensate difference in ability (Becker and Tomes, 1976). Inequality-averse parents allocate their resources equally in sons' and daughters' education whereas more resources may be channelled to the children with higher returns to education if parents focus on efficient use of their resources (Behrman et al., 1982). The latter type of parental characteristic produces gender gaps in education if returns to education vary across gender while the former type is likely to maintain a gender balance. However, some authors have criticized this approach on the ground that parents (father and mother) could rarely have similar preferences over their children's education (Pasqua, 2005). Empirical evidence reveal that parents have asymmetric preferences towards their children's education and thus unitary models do not represent satisfactory results (Doss, 1996).

Additionally, parental decision on investing in children's education may also depend on the perceived future income particularly for those who rely on children's support in their old age (Yueh, 2001). Therefore, parents may take into account not only the returns that accrue to their children but also expected future earnings that children might transfer into household while making the investment decision. Defining personal rate of return to the former and familial return to the latter, this author argues that educational investment among offspring is not influenced by taste or preference for son but due to the efficient use of resources which accounts for both types of returns. Parents will invest more on the children associated with higher returns to education if they take into account the personal return of children whereas they invest on those likely to transfer larger amount of income into the household if familial return is considered.

Whether such parental characteristics produce gender gaps in education is not clear. Reason is that household income comprises of children's labour income, spouses income and other non-labour income. Therefore, parental investment deci-

sion depends on whether to take into account of children's future income alone or total household income. Parents may invest less in girl's education as they move to the husband's house after marriage and are likely to transfer less income into the household relative to sons if only children's own income is considered. However, if likelihood of future transfers from all offspring is same under the assumption of assortative mating parent will invest more in girl's education because this will generate larger household income than the income would be transferred by son.

A bargaining model developed by Pasqua (2005) demonstrates that parental preferences over children's education depend on their respective bargaining power. This approach attempts to extend previous literature that characterise household's decision making process as a single utility maximization problem (see Becker,1974; Becker and Tomes,1976) by allowing parents to have different preferences over consumption, leisure and children's education. Therefore, household decision in investing either in son's or in daughter's education in this model is determined by maximising both parents utility who are allowed to be different in own income level as well as in preferences.

In a another model, the household production model, Yang and Zhu (2003), show that intrafamily allocation of resources depends on parental expectation of their children's roles in household production activities. According to this theoretical framework each family gets involve in household production comprises of two activities namely (a) managerial and (b) execution of production task. Every child is expected to involve in one of these activities. This division of work is central in this analysis to see intrahousehold allocation of schooling. Since education positively affects the quality of decision and managerial decision plays a long lasting effect on family business parents will have incentive to invest more in the children designated as manager than other children who are expected to be a regular worker. This model

demonstrates that families under household production allocate their schooling more unequally than that of families participating in the labour markets. These authors acknowledge the applicability of this theoretical approach especially in developing world where people earn part of their livelihood by working in family-run business.

This model does not explicitly indicate whether parent choose male child to be manager in household production and thus invest more in son's schooling. However, it can be observed that females are not equally given opportunities as males to take a charge of organizational leadership especially in the developing countries. Therefore, this model indicates that anti-bias in female education can exits even in a more extreme form in developing world where the rate of labour force participation is relatively low and the majority of families are involved in family-run enterprises.

2.2.4 Empirical studies

Existing empirical studies on gender gaps in education can be divided into two strands. The first focuses on household characteristics, such as economic status and parental education, and how these influence gender preferences in schooling. In general these studies find that poverty, lack of social security, credit markets and low levels of parental education all contribute to gender biases in educating children. For example, Jacoby and Skoufias (1997) found that rural households in south India vary children's school attendance as self-insurance against seasonal income shocks. While their study did not look for gender differences, when Cameron and Worswick (2001) used the Indonesian Family Survey (1993) data to investigate the same phenomenon from the point of view of gender, they reported a significant bias against females' school attendance. Similarly, using a Pakistani panel data collected by the International Food Policy Research Institute (IFPRI), Sawada (1997) found that households might sacrifice their daughters' education as a coping strategy in the

face of income risk.

The second strand of empirical studies takes into account gender differences in labour market outcomes. Starting from the seminal work of Rosenzweig and Schultz (1982) in rural India, researchers have investigated the link between gender-specific allocation of household resources and gender-specific indicators of labour market outcomes, both indirect ones such as gender-specific survival ratios and direct ones such as returns to education. For instance, using 1995 urban household survey data from Uttar Pradesh, India, Kingdom (1998) found that girls face lower economic rate of returns to education. This is not a robust finding, however. For example, Aslam (2009) decomposes the gender-specific returns to education in the Pakistani labour market between a labour market effect which captures discrimination on the part of employers and a pure education effect. He found that while the pure returns to education were significantly higher for females, the overall market returns were higher for males.

Other authors have also found returns to education to be higher for females in Indonesia (Behram and Deolalikar, 1995) and Bangladesh (Asadullah, 2006). Moreover, Munshi and Rosenzweig (2006) argues that when a traditional institution such as the Indian caste system interacts with modern labour market institutions, the results can favour female education among the lower castes, since males belonging to these castes are more likely to face occupational constraints imposed by the caste system. Using panel data from Bombay over the period 1982-2002, they found that girls were more likely to be enrolled in English-medium schools that prepare students for work in white-collar jobs, while boys were more likely to be enrolled in schools where a local language was used as the medium of instruction and which ultimately led to their students following a traditional occupational path.

To summarize, both the theoretical frameworks and empirical studies do not

clearly indicate the underlying factors of persistence of gender gaps in education. In other words, it is not clear whether low female education can be explained completely by lower expected returns to female education and pro-son parental biases in educational investment.

2.2.5 Marital prospects: Gains from marriage

An important unexplored dimension of the gender gaps in education is marriage. A seminal work to introduce marital prospects into a framework of economic analysis was "A theory of marriage" by Becker (1973, 1974). This framework demonstrates that the gain from marriage relative to remaining single positively depends on their income, human capital and relative difference in wage rates. It shows that couple will enter into marriage if the utility of getting married outweighs the utility from remaining single. The intuition is that married person achieves higher level of utility than being single by utilizing division of labour in household production. For instance, the spouse with a relatively higher labour market outcome will specialize in labour markets whereas the spouse with lower wage can specialize in household works. In turn both of them better off in a marital union. Additionally, sharing same household and other economic and social resources i.e., heating, cooking, housing etc. also yields economies of scale (Becker, 1973).

Becker's view of specialization is not out of critics, however. His framework does not incorporate the possibility of relative advantage in hiring people for household works that allows both spouses to work in labour markets. Opepenheimer et al. (1997) highlight this issue. Similarly, technological advancement can also marginalize the demand for labour in household works allowing both spouses to work in the labour market. Greenwood et al. (2005) and Matouschek and Rasul (2008) point out that decline in labour market discrimination against women and the introduc-

tion of capital goods such as vacuum cleaner and washing machine can reduce the benefit from specialization of spouses. Nonetheless, these authors have not indicated a complete elimination of specialization and thus Becker's theory of marriage can be considered as an important step to explore a new dimension of household economics that analyses marital prospect in a framework of modern economics.

2.2.5.1 Marriage and female education

Becker's theory of marriage does not pay attention on how parents' decisions regarding children's education can be influenced by factors related to marital prospects. Some other papers have shown that the prospect of marriage alone can bias parents against educating their daughters. Lahiri and Self (2007) argue that parents discount the returns to female education since their daughter's future earnings will flow onto their in-laws' household rather than to the natal home. Jafarey (2011) argues that since the marital division of labour encourages some degree of specialization between married couples, the anticipation that a female will get married can discontinuously lower the value of her education relative to her hypothetical self who remains single. However, both of these studies do not incorporate explicitly the age at marriage into their theoretical models. Additionally, hypothesis posited in these models remained to be tested empirically.

Empirical analysis with reference to age at marriage and female education is scant apart from some works carried out by demographers and sociologists but not economist (See for instance; Dahal et al. ,1993; Singh and Sharma, 1996; Bayisenge, 2010; UNICEF, 2011). These descriptive studies explain impact of early marriage on education including other indirect effects such as poor health of children, higher mortality rate, violation of human right etc. but do not establish a casualty between these two variables.

While marriage does not necessarily mean to stop education for those who enter into the marital union but it does in case of females. Several factors can hinder married women's education. Firstly, girls are less likely to attend school after marriage especially in developing countries (Cochrane, 2009; Singh and Sharma, 1996). Additionally, it is difficult for married women to shoulder double responsibilities as wife or mother and schooling at the same time. Some schools even lack female-friendly environment particularly in developing countries since girls often face bullying and abuse by their teacher and mates which makes them feel unsecured and force them to give up schooling (Bayisenge, 2010). The expected future role can also influence female's educational attainment. For example, in a stereo type traditional society females are expected to engage in household chores where education matters less. This encourages females to start a marital life earlier so that they can enjoy larger span of established marital life rather than continuing own education.

Moreover, higher education narrows range of potential mates for girls considering the fact that educated females expect for educated partner, at least not less than their own level of education. This might eventually delay marriage age. Higher education can also delay female's age at marriage by increasing their ability to regulate fertility as well as giving an exposure to western values and behaviour. A relatively high probability of labour force participation and access to better paid jobs which offsets the economic advantage of getting early married is yet an another key factor in delaying female's age at marriage (Singha and Sharma, 1996).

These indicate a positive corelationship between age at marriage and female education. A descriptive statistics from Nepal depicts a similar picture. For instance, the ethnic groups Yadav and Sarki have the lowest age at marriage of 15 years, which is 2 years less than the national average, and also have an average of 2 less years of schooling. Similarly, the ethnic groups, Brahman, Newar and Chettri have higher

average age at marriage and also a higher level of education. The relationship between these two variables is not symmetric for all ethnic groups (see Table 2.6 for details). This implies that the magnitude of correlation between these two variables can largely be influenced by cultural factors and thus can differ significantly across societies. This aspect is given a very less attention while analysing marital behaviour in empirical works reason being either a difficulty to capture heterogeneous socio-cultural variables in the available dataset or has simply been overlooked.

Despite various studies on marital impact on female's wellbeing no attention has yet been paid in empirical studies to establish causality between age at marriage and female education. There have been three studies so far, at least in my knowledge, which examine the interaction between these two variables. The first study carried out by Brien and Lillard (1994) evaluates the relationship between marital status and probability of continuing schooling utilising data from Peninsular, Malaysia. This study also estimates the effect of education on the probability getting married. The second study by Mench et al. (2005) investigate the contribution of urbanization and the levels of schooling on age at marriage. Looking at evidence from 73 developing countries, these authors found that the expansion of schooling has led to a proportional increase in the age at first marriage for females but did not find a similar result for males. This paper, however, did not consider the effect from age at first marriage to education.

Field and Ambrus (2008) look at the effect of early marriage on female schooling and other adult outcomes in Bangladesh. This paper argues that in impoverished and culturally traditional societies, parents have an incentive to marry their daughters young as a form of protection against economic vulnerability. But even in these societies, the age of menarche imposes a constraint on how early girls can be married. Therefore, the authors use the timing of menarche as an instrument in identifying

the impact of early marriage on female schooling. They find that early marriage does significantly lower female schooling and that each year delay in marriage age would increase female schooling by 0.22 years.

2.2.5.2 Marital anticipation and female education

Although the study carried out by Field and Ambrus (2008) is an important step in isolating the effects of age at marriage on female education, the relevance of menarche as an instrumental variable for age at marriage is limited to social settings in which child marriage is prevalent. While this might be true in Bangladesh, it is not necessarily true even in other South Asian countries. Moreover, while child marriage *directly* hinders a female's time in school by imposing spousal duties and possibly motherhood on her at a young age, marriage can exert a more *indirect* disincentive to female education, which applies even to females who marry post-childhood: that is the market return to female education can be lower for a married female than a single one of the same age and level of education. This is because once a woman gets married, in most cultures her burden of non-market household activities increases. Thus, even if a female gets married after the age at which she would normally have finished a particular level of schooling, the sooner she marries after reaching that age, the less likely she will to attain that level in the first place. Field and Ambrus (2008) as well as Brien and Lillard (1994) do not look at the *indirect* disincentive to female education that marriage exerts via division of labour in marital union.

This study contributes to the analysis of the above effect by developing a theory for jointly explaining both female education and anticipated age at marriage. Building on the model of Jafarey (2011) it shows that the earlier the anticipated age at marriage the less schooling a female will receive, despite the fact that there is no direct time conflict between marriage and education in this model.[4] Therefore,

[4]This model rules out child marriage by assuming that schooling and marriage take place in

this study; first develops a theoretical model for jointly determining age at marriage and female education. Second, it tests the causal effect of age at marriage on female education using data from a household survey in Nepal.

2.2.5.3 Social institutions and age at marriage

Despite the fact that marital anticipation affects female education, casualty between these two variables is not straightforward because of endogeniety. Various cultural factors can influence anticipation of age at marriage. In south Asian countries, some ethnographers have pointed out social institution such as dowry plays an important role in determining female marriage age. The types of dowry prevalent in south Asia is effectively a groom price which is found to be positively correlated with educational attainment and social standing of boys (Das, 2009). A case study of rural south Asia shows dowry amount is not only related to grooms own level of education but also with their household wealth (Dalmia and Lawrence, 2005).

Dowry practicing parents will have less incentive to keep their daughter unmarried for older age since an older girl is likely to have more matured and educated boy as her match that will eventually put upward pressure on dowry amount at the time of marriage. A younger girl, on the other hand, matches with a less qualified boy which commands the less amount of dowry (Dhital, 2012). It shows that the institution of dowry compels parents to get their girls married as younger as possible which lowers female's age at marriage in dowry practicing societies.

Taking into account the impact of social institution such as dowry and other socio-cultural factors this study endogenize the age at marriage by introducing variation in the individual ability to benefit in terms of earning power from schooling

adjacent periods of life. However, by introducing a continuing time scale within each period this study is able to treat both education and age at marriage as continuous variables. Since the data set employed in this study depicts a considerable ratio of child marriage, it also tests this empirical model on the sub-set of females who got married after completing childhood.

and showing that females who are more likely to benefit will postpone marriage relative to their counterparts who are less likely. Since theoretical framework developed in this study suggests that females may select into early marriage on the basis of idiosyncratic and unobservable differences in ability, the least-squares estimates will be potentially biased. Thus two-stage least square regression model is used in estimating the impact of marital anticipation on female education.

Nepal is well suited for this study because it has considerable variation in age at first marriage across ethnic groups and communities. In particular, members of the Maithili community, which is concentrated in the regions bordering India, have been identified by ethnographers as practicing an extreme form of dowry culture which leads them to marry their daughters young. The particular dowry practice, locally known as *Tilak Pratha*, which they practice in an extreme form, is effectively a groom price. In order to find a suitable groom, a Maithili girl requires not just physical attributes and a suitable family background, she also needs to pay *Tilak* money, which increases along with the educational qualification and social standing of the boy (Das, 2009). One reason for their strong adherence to this practice is the geographical and cultural proximity of Maithilis to India.

The practice of *Tilak Pratha* induces Maithili parents to marry their daughter as soon as possible since older girls likely to reinforce larger amount of *Tilak* money relative to younger girls. Survey data used in this study shows a significant variation in marital behaviour as well as educational attainment between Maithili and non-Maithili girls.

Figure 2.1 below shows a comparison of age at marriage between Maithili and the others. It shows that the average age of marriage for Maithili community is 15.78 years, almost 2 years lower than the non-Maithili. Similarly, 65% of Maithili women get married before 16 years of age, whereas the average percentage of marriage for

non-Maithili for this age is 42%.[5] This indicates that dowry culture might have a significant role in lowering females' age at marriage within a Maithili community.

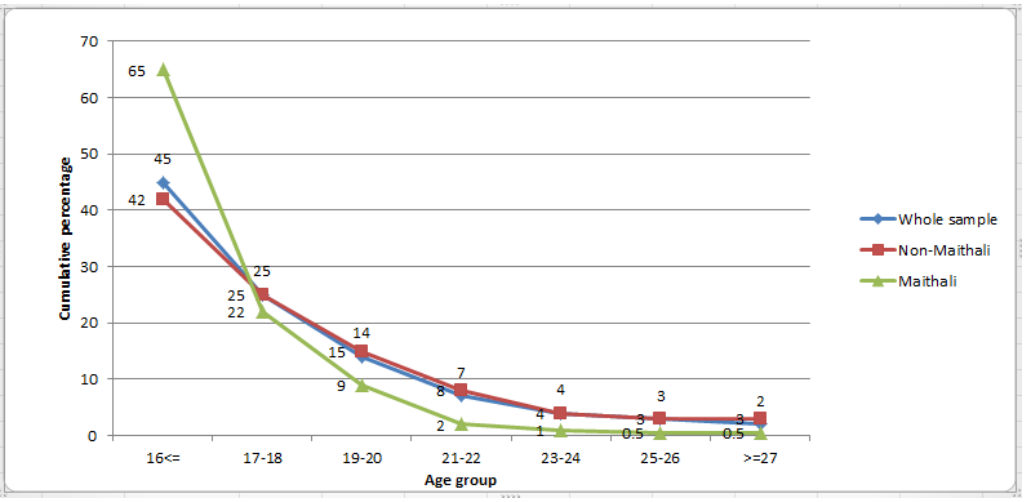

Figure 2.1: Age at marriage comparison

One of the instrumental variables will therefore be a dummy variable indicating membership of the Maithili community. However, since there is the possibility that the Maithili community's cultural attitudes are directly biased against female education (relative to those of other communities), a second instrument, namely the average age at marriage within the respondent's ethno-regional group has been utilized. The second instrument captures joint ethnic and regional variation with regards to marital behaviour and is less likely to be affected by the prevailing cultural attitudes of the respondent's own reference group towards female education. In a two-stage regression it is found that being a member of the Maithili community lowers age at marriage approximately by one year and the effect of this on female education is to lower it by approximately 0.4 years of schooling.

[5]This figures represents from *Sample2549* which corresponds females within the age range of 25-49 years.

2.3 Theoretical framework

This section outlines a framework for jointly determining a female's education and her planned age at marriage. Suppose that a female, indexed by i, is poised to enter adulthood. She has already gone through a period of childhood, in which she has received a level of education, e_i, which for theoretical purposes, is a non-negative, continuous variable. Suppose that her childhood education affects only her adult welfare. In other words, any costs (either explicit or in terms of foregone opportunities) or benefits from education have no effect on her as a child.

Suppose that her time in adulthood is continuous and normalised to the unit interval and that within this interval, she goes through two sub-intervals, single and married. Let $t_i < 1$ be the point of time when she marries, thus it is also the length of time she spends as single and $1 - t_i$ is the length of time she is married.[6]

Following from the above, assume that adult utility can be described by an indirect utility function, V, that varies with each sub-interval of adult life:

$$V_i^j = V(e_i, X_i^j, Z_i),$$

where V_i^j is her utility in each stage j, $j = (s, m)$, s = single and m = married; e_i is her educational level, X_i^j is a set of household, community and environmental characteristics specific to stage j in i's life, and Z_i is a set of factors that are common to both stages in i's adulthood. Z_i could index her ability to benefit from education and convert it into market earnings as well as the innate attitudes of her family and community towards her marriage age and her education. Note that since e_i is determined before reaching adulthood, it is not indexed by j. I assume that V_i^j is increasing in e_i at $e_i = 0$, concave in e_i, and reaches a maximum at some

[6] Alternating stages of matrimonial status as could happen with divorce or widowhood are ruled out.

stage-specific level of education, $\bar{e}_i^j > 0$.

It is also assumed that $\bar{e}_i^s > \bar{e}_i^m$ for all i.[7] The last two assumptions are needed to ensure an interior optimum for e_i.

In addition, it is assumed that there exists a social norm regarding the ideal age of marriage. Let this be denoted by t^*, which applies to all females.[8] Marrying sooner or later than this ideal imposes utility costs in the form of 'loss of face'. There is a long-standing literature in both demography and sociology that have investigated the existence of social and cultural norms regarding age at marriage. While sociologists such as Settersten and Hagestad (1996) and Neugarten et al.(1965) were interested in the broader issue of age norms for various "life-course" transitions, demographers are specifically interested in age at marriage and the length of the reproductive cycle in women (see Billari et al., 2002). Both Neugarten et al. (1965) and Billari et al. (2002) discuss survey evidence on the existence of popular perceptions regarding ideal ages and/or age limits for marriage, the latter from 1960's USA and the former from 1990's Italy. [9]

[7]These assumption are based on Jafarey (2011) where childhood time is explicitly modelled as a choice between developing labour market skills and household skills. While the former requires only schooling time as an input, the latter requires an optimal mix of time in school and time spent at home acquiring domestic training. Too much or too little schooling can result in sub-optimal levels of household skill and this in turn leads to a inverted U-shaped relationship between V^j and e. Since it is assumed, also based on Jafarey(2011), that a female's burden of housework increases after marriage, this leads to the implication that $\bar{e}^s.\bar{e}^m$.

[8]In order to economise on notation, I have left implicit an additional feature of framework which will be used in the empirical part. That is that each female belongs to some reference group, which determines her own ideal age at marriage. But this ideal might vary from group to group.

[9]According to the data cited by Billari et al. (2002), older women perceived age limits for marriage more frequently than younger ones and all age groups believed more strongly in a minimum age than a maximum. For example, 11% of women born between 1945-1947 believed in an upper age limit but only 5% of women born in 1973 did so. These are results from modern Europe. Casual evidence suggests that such culturally influenced age limits are far stronger in traditional South Asian ones than in modern European ones. Unfortunately, similar survey evidence from Asia is not available but even in the 1960s survey data from the USA, Neugarten et al. (1965) reported that 80% of male and 90% of female respondents believed that men should marry between the ages of 20-25 and 85% of male and 90% of women set the analogous age range for women between 19-24.

With this added assumption, her indirect utility can be expressed as

$$V_i = t_i V_i^s(e_i, X_i^s, Z_i) + (1 - t_i)V_i^m(e_i, X_i^m, Z_i) - \delta(t_i - t^*)^2,$$

which is maximised by the appropriate choice of e_i and t_i.

Her educational choice is characterised by the following first order condition (from hereon the agent i subscripts is dropped, unless needed for clarity):

$$t\left[\frac{\partial V^s}{\partial e} - \frac{\partial V^m}{\partial e}\right] + \frac{\partial V^m}{\partial e} = 0,$$

while that for t is

$$V^s - V^m - 2\delta(t - t^*) = 0.$$

Let

$$\Lambda = \frac{\partial V^s}{\partial e} - \frac{\partial V^m}{\partial e}.$$

For the first order condition for e to hold, it must be the case that $\partial V^m/\partial e < 0$ and that $\partial V^s/\partial e > 0$. Thus the optimal level of e lies between $\bar{e}^m$ and $\bar{e}^s$. Note that at the optimal choice of e, $\Lambda > 0$.

To study the mutual dependence of e and t, totally differentiate the first-order condition for e:

$$\left.\frac{\partial e}{\partial t}\right|_t = -\frac{\Lambda}{\Gamma} > 0,$$

where

$$\Gamma = t\frac{\partial^2 V^s}{\partial e^2} + (1 - t)\frac{\partial^2 V^m}{\partial e^2} < 0.$$

Since $\Lambda > 0$ at the point of optimality, e will increase with t.

Turning to the choice of t, it can be solved explicitly from the first-order condi-

tion.

$$t = t^* + \frac{V^s - V^m}{2\delta},$$

which implies that

$$\begin{array}{ccc} > & & > \\ V^s = V^m & \Longrightarrow & t = t^*. \\ < & & < \end{array}$$

In other words, a female delays getting married if her utility from remaining single exceeds utility from being married and expedites marriage otherwise.

By totally differentiating the first-order condition for t,

$$\left.\frac{\partial t}{\partial e}\right|_e = \frac{\Lambda}{2\delta} > 0.$$

Thus t depends positively on e.[10]

2.3.1 Applicability of theory

Theoretical framework developed in this study shows a positive relationship between the anticipated age at marriage and female education in presence of labour market discrimination against women. The main argument is that married female will specialize, to some extent, in household chores in marital union in face of labour market discrimination and thus an anticipation of age at marriage can impose disincentive to acquire education in her childhood. Empirical tests validate this hypothesis.

However, this framework focuses on parental decision on investing in girls' ed-

[10]The above analysis is based on separation of decision making: the educational level is determined taking age at marriage as given; while age at marriage is determined taking education as given. Theoretically an alternative formulation could be to have the educational decision made prior to the age-at-marriage one and taking into account the dependence of the latter on the former. This alternative is unlikely to affect the qualitative predictions of the model and besides it makes more sense in the context of most South Asian countries to assume a separation of authority between mothers, who might exert greater influence on marital decisions and fathers, who might control the allocation of household resources over children's education.

ucation since education takes place in their childhood. Therefore, the parental anticipation of their daughters' age at marriage is shown to have backward effects on females schooling investment. It seems that members of household who allocates household resources (parents) should have control over their children's marital decision to this theory to hold. However, theory can analogously applicable even if female invest on her education by herself as long as she expects to face labour market discrimination. It is because a rational female anticipates specialising in household works in marital union since releasing her spouse in labour market will yields the higher level of utility. This in turn induces female, who anticipate marriage, to invest less on human capital relative to those who expect to remain single.

It shows that theoretical framework developed in this study can be generalized and is expected to explain an ever explored causing factor of anti-bias in female education. However, further test of this hypothesis employing data from different regions and cultural contexts will be an additional advantage to make this theoretical conclusion more robust.

2.4 Data

This paper employs data from the 2003 National Living Standard Survey of Nepal, carried out by its Central Bureau of Statistics with the technical support of the World Bank and UK Department of International Development. The survey follows the World Bank's Living Standard Measurement Survey Strategy and applies a two-step stratified sampling scheme. It took place over 269 Primary Sampling Units, covering 73 out of a total of 75 districts in Nepal and comprises information related to demography, education and literacy, health and maternity, and other information at the household and individual levels. A total of 5240 households and 28110 individuals were included in the sample, and 5028 married females. The data

cover the five administrative regions of Nepal: Eastern, Central, Western, Mid-Western and Far-Western, and an additional category of Abroad for those who were not residing in Nepal at the time of the survey (mostly in India).

The inclusion of all married females to estimate the effect of age at marriage on education may lead to the sample selection bias since unmarried females will be excluded. Table 2.4 reports females' marital status for different age groups. The table shows that the likelihood of marriage increases monotonically until 30 years old at which less than 2.5% will remain unmarried. Therefore, this study considers two sub-samples, $Sample2549$ and $Sample3049$, for the age range of 25-49 and 30-49 years old, in order to consider a sample where potential selection bias because of marriage is minor. These sub-samples thus contain married, divorced, separated and widows in those age ranges. The upper limit of 49 is arbitrarily imposed to exclude potential selection bias because of mortality.

The survey contains two types of educational information on individuals: (1) the highest level of completed schooling, and (2) a categorical question about whether the individual (i) never attended school, (ii) attended in the past and (iii) is currently attending school. Only 28% (from the $Sample2549$ sub-sample) answered question (1). For those respondents who did not answer question (1) but answered question (2-i), their educational level is imputed as zero. This increased the sample size considerably from 1079 to 3760 for $Sample2549$ and from 684 to 2818 for $Sample3049$. The measure of educational achievement derived from question (1) is defined as $Educ1$, and the measure derived by adding to $Educ1$ the imputed values for those who answered question (2-i), as $Educ2$.

The variables used in the econometric analysis are presented in Table 2.7. The average school attainment for married women was 7.51 years using $Educ1$ and fell dramatically to 2.16 years when $Educ2$ is used. Geographically the distribution of

married women was 22%, 34%, 25%, 7% , 4% and 8% from the Eastern, Central, Western, Mid-western, Far-western regions and Abroad, respectively, and 80% live in rural areas. They belong to fifteen different ethnicities.

The upper part of Table 2.8 shows the distribution of marriage age across the sample: 45% were married at or before the age of 16 years. Another 39% were married between 17 and 20. Only 2% of the sample got married after the age of 27 years. There is also a considerably lower age at marriage within the Maithili community as compared to the non-Maithili communities. The lower part of Table 2.8 presents details of the educational background of married females. The majority of married women, 71%, do not appear to have any formal schooling. Of the remainder, only 10% attained primary school, 4% secondary school, 7% high school and 8% received higher education.[11]

2.5 Econometric model and instrumental variables

Establishing a *causal* relationship between female age at marriage and female schooling is not straightforward because of potential endogeneity. In the context of traditional South Asian cultures, there are two potential sources of endogeneity, both arising from the fact that both schooling and marital decisions are effectively in the hands of the girls' parents.

The first is the girl's own ability to benefit from education. Parents invest in a daughter's education according to her expected future labor market earnings, which in turn depends on labor market conditions for female employment and her individual ability to acquire and use human capital. As the theoretical model suggests, if a girl's parent judge her to be of relatively low ability, they may decide both not to school her much, to make better use of her time, and to marry her at an early age

[11]These figures are reported from *Sample2549*.

compared to other girls within her community.

The second is the possibility that in traditional South Asian societies, parents are heavily influenced by social norms that favor early marriage and disfavor schooling of females. Thus social norms could induce a positive relationship between the two variables.

For both reasons, there is a possibility of bi-directional causality between these two variables. Longitudinal data that span enough years could account for such anticipation effects, but are unfortunately not available; only *ex-post* decisions regarding education and the age at marriage can be observed. Given these limitations, OLS estimates of the effects of age at marriage on education are likely to be biased and to be unreliable for this reason.

These issues are addressed using instrumental variables (IVs) to predict a female's age at marriage on the basis of her own, her household's and her community's characteristics. Hypothesis, as reflected in the theoretical model, is that those social norms of the ethnic community to which a female belongs that are important in influencing her marital outcome do not directly influence her education.[12] Indeed, there is evidence from attitudinal surveys that even in communities which have very low rates of female education, all else equal, parents would like to have their daughters receive at least high school-level education (Keiko and Yoshinory, 2006).

This study uses two IVs. The first, taken from ethnographic studies, is the influence of the dowry culture. As stated in the Introduction, the practice of dowry is not only stricter in the Maithili community than in other Nepalese communities, the Maithili custom of linking the value of the dowry to the grooms' economic status en-

[12]This does not preclude community-level norms that also directly affect education, such as a particular community's bias against female education which are controlled for by ethnicity dummies.

courages parents to marry their daughters young. Thus membership of the Maithili community is used as a dummy variable in the age at marriage regression with the expected sign being negative. Note that the survey does not contain information about the dowry paid by the females' parents in the past. Looking at expenses in the current year, Table 2.5 shows that Maithili households report higher dowry and wedding expenses (the latter together with other ceremonies) as a proportion of wealth (proxied by land holding) and income.

However, in using a single community, this instrument could suffer from the potential bias that Maithilis both marry their daughters young and have especially strong unobservable biases against educating their daughters. There is also the possibility of an income effect from large dowries, as argued by Dhital (2012), whereby faced with the choice of paying for their daughters' education or saving up for their dowry, parent choose the latter. By contrast, Dalmia and Lawrence (2005) argue that dowry size is a function of differences in individual and household characteristics between grooms and brides. This suggests that the lower the gap in such characteristics, the smaller will be the dowry payment. This would actually encourage investment in daughters' education. These possibilities have received some attention in the literature and from the limited number of empirical studies on it, the results are mixed.

Dalmia and Lawrence (2005) employ household survey data from the Indian states of Uttar Pradesh and Karnataka, and find that, contrary to their own argument, brides' human capital was positively correlated with the amount of dowry. The authors themselves pointed to two types of possible confounding biases in their data. First, in a polygamous marriage market, a relatively large number of women might have been competing for a limited number of eligible men, and both the educational level of women and the dowry might have reflected this asymmetry between

men and women. Second, both variables might have been positively correlated with household wealth.

Another study carried out by Anderson (2004) estimated the effects of brides' education on dowry payments (parental characteristics and distance to school were used as IVs in the education regression). Employing data from Pakistan this study found a positive relationship between the brides' education and dowry size. However, when the average level of education was controlled for, the estimated coefficient on bride's education became statistically insignificant. These studies make it appear that dowry size might not directly discourage female schooling. Nonetheless in light of this and the possibility of a Maithili-specific bias, this study employs a second IV.

The second instrument is the average age of marriage for the respondent's reference group. This group is defined as the intersection of the ethnic and regional community to which she belongs. Assumption is that the average age of marriage of females in the reference group proxies for the culturally derived ideal age of marriage to which the respondent is expected to aspire. To the extent that there is regional variation in this variable within the same ethnic group, it is hypothesised that this reflects peer-group effects on expected age at marriage but that the innate cultural attitudes towards female education, which might be present within her ethnic group as a whole, have been washed out by this variation.[13] Moreover, a statistical test (Sargan test) carried out in this study has confirmed the exogoneity of IVs.

This study uses three different IVs models: (1) IV1: a Maithili-community binary variable, (2) IV2: average age of females within the ethnic-regional grouping to which

[13]It is assumed that there is no systematic correlation between a particular ethno-regional community's (unobservable) cultural norms towards an appropriate age of female marriage and an appropriate level of girls' schooling, apart from how the former might influence the latter. Note also that, unlike cultural norms regarding age-at-marraige, there is no any literature, at least to my knowledge, which suggests the existence of cultural norms regarding an appropriate level of female schooling, even in cultures that might be generally biased against it.

the respondent belongs, and (3) Two IVs: both. In the first-stage of each regression, age at marriage is regressed on the appropriate IV(s) (and other control variables), and in the second-stage, educational attainment as measured by $Educ1$ and $Educ2$ is regressed on the predicted age at marriage and other control variables.

The two stage regression model can be expressed as,

$$Educ_i = \beta_0 + \beta_1 Mage_i + \beta_2 X_i + u_i, \tag{2.1}$$

$$Mage_i = \delta_0 + \delta_1 Z_i + \delta_2 X_i + v_i, \tag{2.2}$$

where $Educ$ is years of schooling and $Mage$ is age at marriage associated with female i. X comprises a set of exogenous covariates, representing individual as well as household characteristics such as age, father's education, mothers education, household wealth (proxied by price of landholding) and household income. In addition, to capture the possibility of cultural and ethnic variation in valuing female education ethnic dummy variables are also used as explanatory variables. Finally, as there could be a possibility on regional variation on the degree of development which may affect incentive to invest on female education via differences in labour market opportunities, access to credit market as well as supply of educational establishments across regions, urban and regional dummy variables are included in X. See the Appendix 1 for the complete list of variables and their definition. Z is the IV set (IV1, IV2 or Two IVs). u and v are the idiosyncratic error terms associated with female i.

This study is unable to use some of the important exogenous factors in explaining female educational attainment due to the lack of enough information in detecting female's parental households in the data set used.[14] For instance, household size

[14]The survey provides information on parent's identity code and their level of education completed but does not provide household number which constrain to detect parental household of a married female.

and the number of siblings may impose financial constrain on parental decision in investing in their children's education. Similarly, number of female sibling may affect not only on resources available for educational investment but also on her age at marriage. However, household income have been used as explanatory variables as the proxy for the wealth status of parent's household under the assumption of assortative mating.[15]

2.6 Results

2.6.1 Baseline regression results.

Tables 2.10 and 2.12 present the first-stage and Tables 2.11 and 2.13 second-stage baseline regressions, for each sub-sample, respectively. White robust standard errors are reported. Regional and ethnicity dummy variables are included but coefficients not reported.

The first column in Tables 2.11 and 2.13 shows the OLS coefficients. They imply that increasing age at marriage by 1 year is likely to increase female's educational level by .315 years using $Educ1$ and .193 years using $Educ2$ for the sample of 25-49 years-old and .269 years using $Educ1$ and .159 years using $Educ2$ for the sample of 30-49 years-old.

Next consider the three IV models for each subsample where $Mage$ is treated as an endogenous variable. IV1 uses only the Maithili dummy as an instrument; IV2 uses only the average age at marriage by ethnicity and region; and Two IVs uses both instruments together. As expected in IV1, due to the presumed effect of a strong dowry culture, membership to the Maithili community has a significant and negative relationship with age at marriage. IV2 models show a positive effect

[15]Assortative mating implies for homotypic preferences (preferences for self's type) in marriage market.

of average regional and ethnic average marriage age on individual marriage age. All these coefficients are significant at the 1% level. The calculated F-statistics are no less than 11 for the largest 25-49 years old sample indicating strong joint significance of the estimated coefficients, but the same statistic is smaller when the smaller sub-sample of 30-49 years old is used, epecially in $Educ1$ sample. When both IVs are used, the Sargan over-identifying restriction test p-values cannot reject the null hypothesis of validity of the IVs.

Turning to the second-stage results, IVs models shows in general a positive impact of delaying age at marriage ($Mage$) on education although the level of significance varies across models and samples. In most cases, the IV estimates are larger than the OLS ones. This is consistent with the theoretical model and the expectation that ability is positively correlated with both age at marriage and education. That is, parents might delay their daughter's marriage if she has high ability. This could be either by extending the search time for a suitable groom or by benefiting longer from her contribution to their own household income. The standard errors also increased considerably, which determines less precise estimates and thus greater variance in significance levels. For the larger sample of 25-49 years old, IV1 has a coefficient of $Mage$ of .335 (not statistically significant), IV2 .456 and Two Ivs .411 (both significant at 5%) in $Educ1$, and .564, .324, and .361 (significant at 1%), respectively, for $Educ2$.

For the smaller sample of 30-49 years old, $Educ1$ models have no statistically significant results, while for $Educ2$ the coefficient estimates are .346 for IV1 (significant at 5%), .175 for IV2 (significant at 10%) and .212 (significant at 5%). Regarding the other explanatory variables, a quick overview shows that the results are more or less as expected. In the first-stage, father's education, urban and land holding increase age at marriage whereas in the second-stage father's education and urban

increases education. Although not reported, regional and ethnicity dummies have a significant effect on the individual age of marriage and education.

2.6.2 Robustness and validity of the estimated results

This sub-section outlines potential sources of bias in baseline results and attempts to address them. Results are present for the sub-sample of 25-49 years old females only.

First, there is the possibility that poverty drives parents both to keep their daughters out of school and to marry them young so that the burden of maintenance falls on their husbands and in-laws. Moreover, poor parents could be more susceptible to trading off girls' education for the sake of accumulating a sufficient dowry, even in communities that do not practice dowry culture as strictly as Maithilis do.

Second, a potential detrimental effect of early marriage on female education may arise because, unlike our theoretical model in which marriage happens only after the age of schooling has passed, a significant proportion of Nepalese girls get married during childhood and could therefore be obliged to abandon schooling and take up household duties. Both of these sources of bias could affect both our IVs.

A third possible source of bias, affecting only the Maithili instrument, is that this might reflect regional variations in marriage practice, especially regarding age at marriage, rather than an effect of dowry culture specific to Maithilis. This possibility arises because Maithilis are concentrated in certain regions of Nepal that border India; to be precise in four of the six regions of survey data used: Eastern, Central, Western and Abroad. Since cultural practices in Nepal do vary by region and the concomitant degree of urbanization, this could arise as a source of bias.

First, consider the possibility that poverty underlies the observed relationship between female education and age at marriage. This issue is addressed by running

regressions on a restricted sub-sample of households that belong to the upper half of the wealth distribution (proxied by land holding). The results appear in Tables 2.14 and 2.15. Results show that the first-stage coefficients of the instruments increased in value and remain significant at the 1% level across all IV models in $Educ1$ sample. In the second-stage regression the coefficient of age at marriage increased for IV1 and dropped in value and/or significance for IV2 and three IV models. For $Educ2$, a marginal decrease is observed in the second-stage regression in the three IV models. The p-value of the Sargan tests and the F-statistics confirm the validity of the instruments in this sub-sample. The important point is that by and large the qualitative results continue to hold at similar levels of significance, especially in the larger sample.

Secondly, there is the possibility that the detrimental effect of early marriage on female education arises because of child marriages. To filter out this effect, if any, two regressions estimates on sub-samples of females who got married after reaching the age of 14 and 15 have been carried out. The reasons for these cutoff ages are, respectively, 15 is the age set by the International Labor Organization (ILO) convention as the minimum age of employment and one reason for this is that it is the age by which most children will have completed secondary school, while 16 is age at which childhood ends according to Nepal's Children Act, 1992. If child marriage is the main driving force behind low female education it can be expected insignificant effects of age at marriage on education in these sub-samples. The results for the first sub-sample appear in Tables 2.16 and 2.17. Overall the results are similar in magnitude and significance to the corresponding baseline regression models, showing a positive effect of age of marriage on education. The Sargan test rejects the exogeneity of IVs in the $Educ2$ case. Analogous results are observed in the second sub-sample (female married at the age of 16 and above) despite the fact

that age at marriage coefficient in the $Educ1$ sample is statistically significant only at 10% level while it was significant at 5% level in the baseline regression (see Tables 2.18 and 2.19).

The third potential source of bias is that the Maithili instrument might reflect regional variations in marriage practice rather than the effect of Maithili dowry culture. Maithilis are concentrated in regions of Nepal that border India and it is known that cultural practices in Nepal vary by region and the concomitant degree of urbanization. This possibility is addressed by estimating similar models on a sub-sample that comes from regions in which the Maithili community are concentrated. This sub-sample includes the Eastern, Central, Western and Abroad regions but excludes the Mid-Western and Far Western regions. The results appear in Tables 2.20 and 2.21. The estimates reported in Table 2.21 are similar to the corresponding baseline models.

2.7 Conclusion

This study investigated the impact of planned age at marriage on female education on the basis of a theoretical framework for jointly determining both variables which is then has been tested using household data from Nepal. In light of the framework, instruments are developed that could control for the potential endogeneity of main explanatory variable and then employed an instrumental variables procedure for identifying its impact on female education.

Results suggest that a strict adherence to dowry practices, as in the Maithili community, lowers age at marriage while the average age of marriage of one's ethno-regional group increases it. It is found that marital behaviours that favour early marriage significantly reduce female educational attainment. While the results differed slightly across different instrumental variables and samples, a consensus es-

timate would be that each year's delay in marriage increases female education by 0.4 years. This figure is roughly in line with but slightly higher than that found by Field and Ambrus (2008) for Bangladesh.

All models on sub-samples of the data are tested in order to control for potential bias. These were the possibilities that (i) the positive association of female education with age at marriage could reflect the results of a coping mechanism amongst the poorest households; (ii) the high incidence of child marriage in Nepal could have induced estimated coefficients through a more direct *ex post* mechanism rather than the more indirect *ex ante* mechanism stressed by the theory; (iii) the concentration of Maithilis in certain regions of Nepal could have led to results which reflect regional variations rather than the dowry culture of Maithilis. Overall robustness results continue to suggest a negative impact of early marriage on female education.

Table 2.1: Regional literacy rate (%) by gender

Year	*1980*			*1995*		*2000*	
	Male	Female		Male	Female	Male	Female
World	77.2	61.9		83.6	71.2	85.2	73.6
Developing countries	68.9	46.8		78.9	61.7	81.2	65.5
Sub-Saharan Africa	51.8	29.2		66.6	47.3	70.9	50.3
Arab state	55	26.2		68.4	44.2	72.2	50.1
Latin America	82.1	77.5		87.7	85.5	89	87.4
Eastern Asia/Ociana	80.4	58		90.6	76.3	92.8	80.6
Southern Asia	52.8	24.5		62.9	36.6	66.6	40.7
Least Developed countries	48.3		24.9	59.5	38.1	62.9	42.4
Developed countries	98	95.4		98.9	98.4	99.1	98.8

Source: UNESCO, cited from Yakozeki (2004).

Table 2.2: Percentage of female education by educational level and by region

Educational level	Primary				Secondary				Tertiary			
Year	1960	1980	2000	2025	1960	1980	2000	2025	1960	1980	2000	2025
World	43	45	47	47	41	43	45	46	34	43	44	44
Developing	39	44	46	47	29	39	44	46	26	35	40	41
Sub-Saharan	34	43	44	45	25	34	40	40	11	21	28	27
Arab States	34	41	45	46	26	37	45	46	17	31	40	42
Latin America	48	49	49	49	47	50	51	50	30	43	47	47
Eastern Asia	39	45	48	50	30	40	47	51	24	24	32	32
Southern Asia	36	42	45	46	25	36	41	44	26	33	39	41
LDCs	32	40	44	44	18	31	39	39	16	26	31	31

Source: UNESCO, cited from Yakozeki (2004).

Table 2.3: Gender Parity Index (GPI) on educational enrollment in south Asia: 2002

Country/region	Pre-primary	Primary	Secondary	Upper secondary
World	.99	.93	.93	.92
SAARC	.98	.86	.82	.76
Bangladesh	1	-	1.21	.93
Bhutan	-	.92	92	.74
India	1.03	.87	.78.69	-
Maldives	.99	.96	1.17	.88
Nepal	.85	.86	.78	.68
Pakistan	-	.68	-	-
Sri Lanka	-	.99**	1.01**	1.14

Source: Institute for statistics, UNESCO.
Note: - Indicates data not available and ** indicates GPI based on previous year. SAARC represents eight countries in south-Asia: Sri Lanka, Bhutan, India, Maldives, Nepal, Pakistan, Bangladesh and Afghanistan.

Table 2.4: Females' marital status by age group (in %)

Age group	Married	Divorced	Separated	Widow	Unmarried
<= 15	2.76	0.10	-	0.05	97.09
16-20	42.98	0.19	0.51	0.13	56.20
21-24	76.08	0.10	0.52	0.21	23.09
25-29	90.76	-	0.84	0.65	7.74
30-34	93.74	-	1.76	2.09	2.41
35-39	93.35	0.18	0.72	3.60	2.16
40-44	89.49	0.26	2.37	5.78	2.10
45-49	83.22	0.34	2.37	12.37	1.69
Total*	59.51	0.14	1.15	8.91	30.29

Notes: * all ages, including age> 49.

Table 2.5: Wedding expenses: Current year (in '000)

	Maithili			*Non-Maithili*		
	Total	*Rural*	*Urban*	*Total*	*Rural*	*Urban*
D-cost	19.39(38.13)	20.84(39.92)	7.04(11.86)	5.66(26.70)	5.55(28.7)	5.99(21.90)
Obs.	57	48	9	764	582	182
Wed-exp	7.62(20.96)	7.75(21.41)	5.38(11.10)	8.41(32.37)	6.43(16.15)	19.31(72.50)
Obs.	232	220	12	1776	1503	273
D-cost/Lnholding	.219	.251	0.060	.020	.029	.015
D-cost/Hincome	.005	.015	.0004	.0005	.0005	.0005
Wed-exp/Lnholding	.088	.092	.041	.044	.040	.054
Wed-exp/Hincome	.004	.004	.003	.001	.001	.0007

Note: D-cost=dowry paid, Wed-exp= marriage, birth and other ceremonies expenses. These figures represent aggregate household data. Standard deviations in parentheses.

Table 2.6: Descriptive statistics: 1

Ethnicity	*Mage*		*Educ*		*Wedding cost ('000)*		*Dowry cost ('000)*	
	Rural	*Urban*	*Rural*	*Urban*	*Rural*	*Urban*	*Rural*	*Urban*
Brahmin	17.28(3.08)	20.28(4.89)	7.72(3.20)	10.57(2.96)	10.85(30.43)	5.71(7.75)	6.57(26.55)	6.98(24.79)
Chetri	17.49(3.08)	18.44(3.47)	6.59(3.12)	8.30(3.22)	5.94(14.21)	8.64(12.32)	3.47(12.93)	1.89(2.74)
Newar	18.67(3.22)	20.50(3.60)	7.45(3.59)	9.21(3.45)	12.53(22.48)	11.54(26.33)	8.62(49.45)	8.09(27.30)
Magar	17.51(2.88)	19.33(0.77)	5.77(2.73)	7.40(3.36)	3.66(5.72)	1.45(1.48)	1.63(5.21)	2.90(2.28)
Tharu	16.17(2.59)	16.50(2.12)	6.26(3.00)	4.00(2.82)	2.81(5.03)	0.00(0.00)	2.31(5.24)	0.80(0.00)
Tamang	17.60(3.11)	17.55(1.94)	5.09(3.12)	7.80(3.35)	3.52(4.01)	2.25(8.02)	1.49(1.61)	2.43(1.79)
Kami	16.79(3.05)	19.6(4.39)	5.34(2.93)	5.42(2.76)	2.73(4.51)	6.00(4.24)	2.46(6.09)	2.33(2.30)
Yadav	15.01(2.62)	17.33(1.52)	6.03(2.57)	0.00(0.00)	3.18(5.67)	0.00(0.00)	12.15(26.67)	0.00(0.00)
Muslim	15.21(2.52)	16.71(2.26)	4.39(2.89)	6.00(2.16)	5.91(11.49)	6.62(9.13)	14.48(23.16)	13.04(24.67)
Rai	19.36(3.80)	20.57(3.15)	5.95(3.14)	7.50(3.41)	6.18(14.56)	0.00(0.00)	2.11(2.31)	.50(0.00)
Gurung	19.05(4.13)	20.50(2.94)	6.28(3.42)	8.80(2.77)	3.01(4.75)	3.53(2.54)	1.61(1.67)	1.10(1.14)
Damai	17.01(3.55)	18.33(1.52)	5.44(2.52)	6.60(1.03)	5.13(8.97)	22.16(43.57)	8.45(15.15)	0.00(0.00)
Limbu	20.59(5.18)	0.00(0.00)	6.22(2.86)	0.00(0.00)	19.68(24.42)	0.00(0.00)	0.40(0.00)	0.00(0.00)
Sarki	15.66(2.79)	14.00(0.00)	4.83(2.71)	0.00(0.00)	3.02(5.00)	2.75(1.06)	1.35(1.80)	0.00(0.00)
Otherethnic	16.26(2.96)	17.71(3.83)	6.16(3.30)	8.09(3.35)	6.46(18.23)	12.18(19.47)	12.93(43.49)	9.72(34.15)

-Standard errors in parentheses.

Table 2.7: Descriptive statistics: 2

Variable	*Sample2549*	*Sample3049*
*Educ*1	7.51(.106)	7.37(.136)
*Educ*2	2.16(.063)	1.78(.068)
Feduc	2.75(.138)	2.75(.175)
Meduc	.649(.070)	.611(.087)
Urban	.202(.012)	.219(.015)
Lnholding('00000)	3.008(.469)	3.11(.536)
Hhincome('00000)	177.54(41.98)	141.93(38.88)
Mage	19.00(.107)	18.97(.144)
Age dummies		
25-29	.366(.014)	-
30-34	.217(.012)	.342(.018)
35-49	.220(.012)	.346(.018)
40-44	.108(.008)	.178(.014)
45-49	.084(.008)	.133(.012)
Ethnic dummies		
Brahman	.255(.013)	.271(.017)
Chettri	.155(.011)	.160(.014)
Newar	.253(.013)	.276(.017)
Magar	.046(.006)	.042(.007)
Tharu	.022(.004)	.017(.005)
Tamang	.021(.004)	.014(.004)
Kami	.012(.003)	.010(.003)
Yadav	.010(.003)	.005(.002)
Muslim	.012(.003)	.010(.003)
Rai	.024(.004)	.016(.004)
Gurung	.032(.005)	.035(.007)
Damai	.010(.003)	.004(.002)
Limbu	.012(.003)	.010(.003)
Sarki	.001(.001)	-
Others	.142(.010)	.124(.012)
Regional dummies		
Eastern	.216(.012)	.192(.015)
Central	.335(.014)	.365(.018)
Western	.253(.013)	.248(.016)
Mid-western	.072(.007)	.077(.010)
Far-western	.033(.005)	.030(.006)
Abroad	.076(.008)	.071(.009)

Notes: Standard deviations in parentheses.

Table 2.8: Comparison of age at marriage and educational distribution

	Sample2549			*Sample3049*		
Variable	All	Non-Maithili	Maithili	All	Non-Maithili	Maithili
			Age at marriage			
Average	17.44(.059)	17.71(.064)	15.78(.127)	17.34(.078)	17.61(.077)	15.48(.157)
Married ≤16	.45(.008)	.42(.008)	.65(.021)	.47(.009)	.44(.009)	.69(.024)
Married 17-18	.25(.007)	.25(.007)	.22(.018)	.24(.008)	.25(.008)	.20(.021)
Married 19-20	.14(.005)	.15(.006)	.09(.012)	.14(.006)	.15(.007)	.08(.014)
Married 21-22	.07(.004)	.08(.004)	.02(.006)	.07(.004)	.07(.005)	.02(.006)
Married 23-24	.04(.003)	.04(.003)	.01(.0050	.03(.003)	.03(.003)	.005(.002)
Married 25-26	.03(.002)	.03(.003)	.005(.002)	.03(.003)	.03(.003)	.005(.002)
Married ≥27	.02(.002)	.02(.002)	.005(.002)	.02(.002)	.03(.003)	.005(.002)
			Education			
No formal schooling	.71(.007)	.69(.008)	.87(.014)	.76(.008)	.73(.008)	.92(.014)
Primary [1-5]	.10(.004)	.11(.005)	.06(.010)	.09(.005)	.10(.005)	.04(.009)
Secondary [6-7]	.04(.003)	.03(.003)	.03(.007)	.03(.003)	.03(.003)	.01(.006)
High school [8-10]	.07(.004)	.08(.004)	.02(.006)	.06(.004)	.07(.005)	.01(.005)
Higher education [≥11]	.08(.004)	.09(.004)	.02(.006)	.06(.004)	.07(.007)	.02(.007)
Obs.	3760	3244	516	2818	2460	358

Notes: Grades corresponding to each educational level from variable *Educ*2 are presented in brackets. Standard deviations in parentheses.

Table 2.9: Distribution of the Maithili community

Ethnicity	*Sample2549*	*Sample3049*
	Ethnic distribution	
Bramhin	.02(.005)	.02(.007)
Yadav	.16(.016)	.17(.019)
Muslim	.14(.014)	.13(.017)
Sarki	.03(.007)	.02(.007)
Tharu	.04(.008)	.03(.009)
Other	.61(.018)	.63(.025)
	Regional distribution	
Eastern	.38(.021)	.38(.025)
Central	.36(.021)	.36(.025)
Western	.01(.003)	.01(.004)
Mid-western	-	-
Far-western	-	-
Abroad(India)	.25(.018)	.25(.015)
Obs.	516	358

Notes: Standard deviations in parentheses.

Table 2.10: First-stage regression results: Baseline model (*Sample2549*)

Variable	**IV1**	**IV2**	**Two IVs**
	Dep.var. Mage, sub-sample for *Educ1*		
Avmage	-	.737***(.167)	.705***(.167)
Maithili	-1.70***(.503)	-	-1.58***(.500)
Age:30-34	-.169(.272)	-.083(.270)	-.164(.270)
Age:35-39	-.295(.269)	-.259(.268)	-.305(.267)
Age:40-44	-.135(.341)	-.115(.339)	-.145(.338)
Age:45-49	-.722*(.381)	-.683*(.379)	-.747*(.378)
Feduc	.094***(.025)	.090***(.025)	.094***(.025)
Meduc	-.046(.049)	-.034(.049)	-.039(.049)
Urban	1.22***(.311)	1.09***(.312)	1.05***(.311)
Lnholding	1.33*(.792)	1.28*(.788)	1.26*(.783)
Hincome	.011(.009)	.011(.009)	.011(.009)
Obs.	1079	1079	1079
R^2	.1644	.1704	.1775
	Dep.var. Mage, sub-sample for *Educ2*		
Avmage	-	.997***(.099)	.976***(.099)
Maithili	-.890***(.200)	-	-.790***(.198)
Age:30-34	-.083(.165)	-.108(.163)	-.127(.163)
Age:35-39	-.235*(.154)	-.253*(.152)	-.278*(.152)
Age:40-44	-.439***(.165)	-.446***(.163)	-.478***(.163)
Age:45-49	-.818**(.182)	-.847***(.180)	-.861***(.180)
Feduc	.115***(.021)	.114***(.020)	.115***(.020)
Meduc	-.006(.046)	.002(.045)	-.003(.045)
Urban	1.64***(.238)	1.44***(.237)	1.42***(.236)
Lnholding	1.33**(.673)	1.26*(.669)	1.24*(.665)
Hincome	.014(.011)	.013(.011)	.013(.010)
Obs.	3760	3760	3760
R^2	.1921	.2031	.2063

Notes: Robust standard errors in parentheses. *** significant at 1%, ** significant 5%, * significant at 10% level. IV1: Maithili. IV2: Avmage. Age:25-29 as base category. Regional and ethnicity dummies are included but not reported.

Table 2.11: Second-stage regression results: Baseline model (*Sample2549*)

Variable	**OLS**	**IV1**	**IV2**	**Two IVs**
		Dep.var. *Educ1*		
Mage	.315***(.028)	.335(.254)	.456**(.199)	.411**(.161)
Age:30-34	-.348*(.236)	-.347*(.235)	-.337(.237)	-.340(.236)
Age:35-39	-.437*(.233)	-.432*(.240)	-.402*(.240)	-.413*(.236)
Age:40-44	-.551*(.297)	-.549*(.295)	-.537*(.298)	-.541*(.296)
Age:45-49	-1.19***(.372)	-1.18***(.368)	-1.10***(.357)	-1.13***(.346)
Feduc	.184***(.020)	.183***(.031)	.172***(.028)	.176***(.026)
Meduc	.091***(.031)	.092**(.044)	.097**(.044)	.095**(.043)
Urban	1.86***(.279)	1.83***(.419)	1.68***(.371)	1.73***(.338)
Lnholding	-.534(.679)	-.563(.729)	-.730(.647)	-.668(.641)
Hincome	.001(.003)	.001(.003)	.001(.003)	.001(.003)
IVs F-statistic		[11.52]	[19.30]	[14.72]
Sargan test p-value				{.7023}
R^2	.3625	.3621	.3458	.3546
		Dep.var. *Educ2*		
Mage	.193***(.017)	.564***(.213)	.324***(.088)	.361***(.083)
Age:30-34	-.649***(.156)	-.627***(.157)	-.641***(.146)	-.639***(.147)
Age:35-39	-1.14***(.141)	-1.07***(.152)	-1.12***(.137)	-1.11***(.137)
Age:40-44	-1.61***(.145)	-1.46***(.179)	-1.56***(.150)	-1.54***(.151)
Age:45-49	-1.64***(.156)	-1.34***(.243)	-1.54***(.176)	-1.51***(.175)
Feduc	.385***(.023)	.343***(.031)	.371***(.021)	.366***(.020)
Meduc	.211***(.041)	.211***(.043)	.211***(.040)	.211***(.040)
Urban	2.64***(.291)	2.02***(.424)	2.42***(.258)	2.36***(.254)
Lnholding	1.06(.807)	.560(.713)	.886(.758)	.836(.742)
Hincome	.010(.010)	.005(.007)	.008(.009)	.008(.009)
IVs F-statistic		[19.76]	[101.17]	[26.43]
Sargan test p-value				{.1419}
R^2	.4314	.2105	.2224	.4282

Notes: Robust standard errors in parentheses. *** significant at 1%, ** significant 5%, * significant at 10% level. Mage is treated as endogenous. IV1: Maithili. IV2: Avmage. Age:25-29 as base category. Regional and ethnicity dummies are included but not reported.

Table 2.12: First-stage regression results: Baseline model (*Sample3049*)

Variable	**IV1**	**IV2**	**Two IVs**
	Dep.var. Mage, sub-sample for *Educ1*		
Avmage	-	.822***(.203)	.785***(.207)
Maithili	-1.60*(.803)	-	-1.30*(.794)
Age:35-39	-.151(.327)	-.191(.322)	-.162(.323)
Age:40-44	.021(.396)	-.048(.392)	.001(.397)
Age:45-49	-.545(.424)	-.583(.411)	-.567(.415)
Feduc	.108***(.031)	.106***(.031)	.108***(.031)
Meduc	-.063(.065)	-.045(.064)	-.048(.064)
Urban	1.51***(.447)	1.39***(.445)	1.37***(.450)
Lnholding	.720(1.01)	.672(.999)	.651(.995)
Hincome	.029***(.003)	.028***(.003)	.027***(.003)
Obs.	684	684	684
R^2	.1650	.2113	.2148
	Dep.var. Mage, sub-sample for *Educ2*		
Avmage	-	1.00***(.111)	.974***(.111)
Maithili	-1.13***(.221)	-	-1.01***(.219)
Age:35-39	-.145(.162)	-.136(.161)	-.144(.160)
Age:40-44	-.370**(.183)	-.345*(.181)	-.362**(.180)
Age:45-49	-.750***(.209)	-.751***(.207)	-.746***(.206)
Feduc	.112***(.026)	.113***(.026)	.113***(.026)
Meduc	-.024(.060)	-.011(.059)	-.017(.060)
Urban	1.96***(.352)	1.78***(.356)	1.75***(.356)
Lnholding	.719(.789)	.649(.785)	.623(.774)
Hincome	.036***(.003)	.035***(.004)	.035***(.003)
Obs.	2818	2818	2818
R^2	.1956	.2104	.2071

Notes: Robust standard errors in parentheses. *** significant at 1%, ** significant 5%, * significant at 10% level. IV1: Maithili. IV2: Avmage. Age:30-34 as base category. Regional and ethnicity dummies are included but not reported.

Table 2.13: Second-stage regression results: Baseline model (*Sample3049*)

Variable	**OLS**	**IV1**	**IV2**	**Two IVs**
		Dep.var. *Educ1*		
Mage	.269***(.034)	-.054(.446)	.169(.169)	.125(.161)
Age:35-39	-.140(.268)	-.200(.294)	-.158(.268)	-.167(.269)
Age:40-44	-.264(.324)	-.277(.343)	-.268(.320)	-.270(.323)
Age:45-49	-.990**(.392)	-1.17**(.482)	-1.04***(.401)	-1.07***(.402)
Feduc	.203***(.026)	.238***(.056)	.214***(.031)	.219***(.031)
Meduc	.106***(.040)	.086*(.048)	.100***(.039)	.097***(.039)
Urban	1.81***(.371)	2.31***(.794)	1.97***(.443)	2.03***(.441)
Lnholding	-1.24(.729)	-1.01(1.01)	-1.17*(.794)	-1.14(.825)
Hincome	.006(.005)	.016(.013)	.009(.007)	.011*(.007)
IVs F-statistic		[3.98]	[16.38]	[9.35]
Sargan test p-value				{.6029}
R^2	.2739	.3594		
		Dep.var. *Educ2*		
Mage	.159***(.018)	.346**(.148)	.175*(.098)	.212**(.083)
Age:35-39	-.500***(.149)	-.475***(.152)	-.498***(.149)	-.493***(.149)
Age:40-44	-.964***(.152)	-.898***(.163)	-.959***(.156)	-.946***(.154)
Age:45-49	-1.00***(.162)	.862***(.202)	-.991***(.179)	-.963***(.175)
Feduc	.390***(.029)	.369***(.033)	.388***(.032)	.384***(.031)
Meduc	.233***(.053)	.236***(.052)	.233***(.052)	.234***(.052)
Urban	2.64***(.353)	2.26***(.862)	2.60***(.392)	2.53***(.381)
Lnholding	.936(1.27)	.796(1.19)	.924(1.26)	.897(1.25)
Hincome	.029*(.015)	.023*(.015)	.029*(.015)	.028*(.015)
IVs F-statistic		[26.30]	[80.37]	[51.70]
Sargan test p-value				{.3055}
R^2	.4256	.3939	.4253	.4230

Notes: Robust standard errors in parentheses. *** significant at 1%, ** significant 5%, * significant at 10% level. Mage is treated as endogenous. IV1: Maithili. IV2: Avmage. Age:30-34 as base category. Regional and ethnicity dummies are included but not reported.

Table 2.14: First-stage regression results: Upper wealth households

	Dep.var. Mage, sub-sample for *Educ1*		
Avmage	-	.964***(.181)	.940***(.184)
Maithili	-2.08***(.581)	-	-1.79***(.586)
Age:30-34	-.223(.374)	-.145(.364)	-.218(.361)
Age:35-39	-.633*(.372)	-.528(.367)	-.598*(.365)
Age:40-44	-.544(.501)	-.676(.496)	-.633(.494)
Age:45-49	-.963*(.569)	-1.13**(.536)	-1.21**(.532)
Feduc	.064(.039)	.066*(.038)	.074(.038)
Meduc	.001(.084)	-.006(.083)	.000(.082)
Urban	1.58***(.578)	1.35***(.595)	1.31***(.593)
Lnholding	1.07(.806)	1.00(.795)	.929(.787)
Hincome	-.003(.003)	-.003(.003)	-.003(.003)
Obs.	530	530	530
R^2	.2036	.2271	.2553
	Dep.var. Mage, sub-sample for *Educ2*		
Avmage	-	.960***(.117)	.927***(.119)
Maithili	-1.22***(.260)	-	-1.09***(.258)
Age:30-34	-.139(.212)	-.177(.209)	-.185(.208)
Age:35-39	-.618***(.192)	-.654***(.189)	-.682***(.189)
Age:40-44	-.671***(.239)	-.706***(.236)	-.727***(.202)
Age:45-49	-1.10***(.272)	-1.14***(.268)	-1.16***(.266)
Feduc	.091***(.031)	.097***(.030)	.100***(.030)
Meduc	.043(.074)	.038(.073)	.038(.072)
Urban	1.96***(.604)	1.77***(.013)	1.70***(.613)
Lnholding	1.45**(.738)	1.27*(.720)	1.26*(.713)
Hincome	-.006(.005)	-.006(.005)	-.006(.005)
Obs.	1877	1877	1877
R^2	.1975	.1997	.2052

Notes: Robust standard errors in parentheses. *** significant at 1%, ** significant 5%, * significant at 10% level. IV1: Maithili. IV2: Avmage.
-Age:25-29 as base category.

Table 2.15: Second-stage regression results: Upper wealth households

Variable	**OLS**	**IV1**	**IV2**	**Two IVs**
		Dep.var. *Educ1*		
Mage	.300***(.046)	.500**(.246)	.335**(.157)	.364***(.140)
Age:30-34	-.428(.356)	-.399(.358)	-.435(.347)	-.431(.347)
Age:35-39	-.576*(.339)	-.465(.357)	-.540*(.331)	-.526*(.329)
Age:40-44	-1.03***(.432)	-.916*(.473)	-1.01**(.429)	-.996**(.432)
Age:45-49	-2.02***(.496)	-1.85***(.532)	-2.08***(.492)	-.198***(.492)
Feduc	.174***(.035)	.163***(.036)	.175***(.035)	.173***(.035)
Meduc	.122***(.063)	.124*(.067)	.120*(.062)	.120*(.063)
Urban	1.89***(.553)	1.55**(.699)	1.94***(.577)	1.89***(.569)
Lnholding	-.260(.743)	-.159(.850)	-.316(.724)	-.272(.739)
Hincome	-.005*(.003)	-.005*(.003)	-.005*(.003)	-.005*(.003)
IVs F-statistic	-	[12.90]	[28.31]	[17.31]
Sargan test p-value				{.6935}
R^2	.3622	.3315	.3647	.3628
		Dep.var. *Educ2*		
Mage	.182***(.025)	.384**(.185)	.313***(.118)	.330***(.103)
Age:30-34	-.926***(.238)	-.900***(.239)	-.910***(.236)	-.907***(.236)
Age:35-39	-1.44***(.211)	-1.32***(.235)	-1.36***(.215)	-1.35***(.213)
Age:40-44	-2.09***(.207)	-.196***(.245)	-2.01(.202)	-2.00(.217)
Age:45-49	-2.09***(.216)	-1.88(.298)	-1.95***(.244)	-1.93***(.238)
Feduc	.371***(.037)	.353***(.039)	.359***(.038)	.358***(.037)
Meduc	.272***(.074)	.263***(.074)	.267***(.073)	.266***(.074)
Urban	1.70***(.555)	1.28*(.705)	1.43**(.605)	1.39**(.602)
Lnholding	1.98*(1.02)	1.81*(1.00)	1.75*(.958)	1.76*(.958)
Hincome	-.005(.007)	-.004(.007)	-.004(.006)	-.004(.006)
IVs F-statistic		[22.19]	[66.38]	[40.73]
Sargan test p-value				{.7318}
R^2	.3786	.3464	.3650	.3613

Notes: Robust standard errors in parentheses. *** significant at 1%, ** significant 5%, * significant at 10% level. IV1: Maithili. IV2: Avmage.
-Age:25-29 as base category.

Table 2.16: First-stage regression results : Adult marriage (Mage $\geq$ 15)

Variable	**IV1**	**IV2**	**Two IV**
	Dep.var. Mage, sub-sample for *Educ1*		
Avmage	-	.720***(.171)	.712***(.173)
Maithili	-.895*(.522)	-	-.825*(.524)
Age:30-34	-.001(.259)	.028(.255)	-.006(.258)
Age:35-39	-.145(.269)	-.145(.267)	-.172(.264)
Age:40-44	-.235(.256)	-.244(.356)	-.267(.356)
Age:45-49	-.371(.393)	-.376(.367)	-.410(.372)
Feduc	.078***(.023)	.078***(.023)	.079***(.023)
Meduc	-.016(.045)	-.006(.045)	-.008(.045)
Urban	1.32***(.315)	1.18***(.318)	1.17***(.319)
Lnholding	1.10*(.743)	1.03(.734)	1.02(.732)
Hincome	.009(.009)	.008(.009)	.008(.009)
Obs.	1003	1003	1003
R^2	.1293	.1442	.1456
	Dep.var. Mage, sub-sample for *Educ2*		
Avmage	-	.770***(.095)	.751***(.096)
Maithili	-.752***(.166)	-	-.656***(.165)
Age:30-34	.116(.144)	.098(.141)	.082(.141)
Age:35-39	-.057(.138)	-.072(.138)	-.092(.137)
Age:40-44	.038(.164)	.039(.162)	.009(.162)
Age:45-49	-.032(.200)	-.069(.199)	-.091(.198)
Feduc	.085***(.020)	.086***(.020)	.086***(.019)
Meduc	.014(.043)	.022(.044)	.017(.043)
Urban	1.56***(.269)	1.39***(.272)	1.38***(.271)
Lnholding	1.19*(.705)	1.10*(.699)	1.09*(.696)
Hincome	.010(.011)	.010(.010)	.010(.010)
Obs.	3128	3128	3128
R^2	.1512	.1725	.1672

Notes: Robust standard errors in parentheses. *** significant at 1%, ** significant at 5%, * significant at 10% level. IV1: Maithili. IV2: Avmage.
-Age:25-29 as base category.

Table 2.17: Second-stage regression results: Adult marriage (Mage $\geq$ 15)

Variable	**OLS**	**IV1**	**IV2**	**Two IV**
		Dep.var. *Educ1*		
Mage	.294***(.031)	.365(.486)	.416*(.217)	.410**(.199)
Age:30-34	-.459*(.245)	-.462*(.243)	-.464*(.243)	-.464*(.343)
Age:35-39	-.498**(.242)	-.490*(.245)	-.484*(.243)	-.485**(.242)
Age:40-44	-.632**(.307)	-.617*(.320)	-.606*(.313)	-.607**(.312)
Age:45-49	-1.26***(.410)	-1.24***(.436)	-1.22***(.413)	-1.22***(.411)
Feduc	.187***(.021)	.181***(.043)	.178***(.026)	.178***(.025)
Meduc	.081**(.032)	.082**(.033)	.083**(.033)	.083**(.033)
Urban	1.88***(2.88)	1.79**(.714)	1.72***(.404)	1.73***(.389)
Lnholding	-.606(.603)	-.687(.818)	-.750(.624)	-.742(.620)
Hincome	.001(.003)	.001(.005)	.001(.003)	.001(.002)
IVs F-statistic	-	[2.93]	[17.61]	[10.01]
Sargan test p-value				{.9232}
R^2	.3479	.3441	.3368	.3379
		Dep.var. *Educ2*		
Mage	.198***(.023)	1.01***(.300)	.366***(.129)	.468***(.120)
Age:30-34	-.765***(.175)	-.875***(.210)	-.787***(.176)	-.801***(.178)
Age:35-39	-1.27***(.158)	-1.24***(.189)	1.26***(.158)	-1.26***(.160)
Age:40-44	-1.76***(.168)	-1.82***(.216)	-1.77***(.170)	-1.78***(.174)
Age:45-49	-1.81(.185)	-1.80(.256)	1.81***(.190)	-1.81***(.196)
Feduc	.400(.025)	.330***(.024)	.385***(.027)	.377(.026)
Meduc	.190***(.046)	.174***(.055)	.187***(.045)	.185***(.046)
Urban	2.71***(.310)	1.42**(.593)	2.44***(.367)	2.28***(.360)
Lnholding	1.12(.870)	.139(.771)	.922(.814)	.798(.777)
Hincome	.008(.010)	.001(.006)	.007(.008)	.006(.008)
IVs F-statistic		[20.45]	[64.39]	[38.99]
Sargan test p-value				{.0193}
R^2	.4465	.0997	.4318	.4086

Notes: Robust standard errors in parentheses. *** significant at 1%, ** significant at 5%, * significant at 10% level. IV1: Maithili. IV2: Avmage.
-Age:25-29 as base category.

Table 2.18: First-stage regression results : Adult marriage (Mage $\geq$ 16)

Variable	**IV1**	**IV2**	**Two IV**
	Dep.var. Mage, sub-sample for *Educ1*		
Avmage	-	.707***(.164)	.704***(.164)
Maithili	-.406(.541)	-	-.322(.542)
Age:30-34	.328(.259)	.343(.253)	.327(.257)
Age:35-39	-.044(.272)	-.064(.271)	-.075(.270)
Age:40-44	-.173(.356)	-.192(.356)	-.201(.356)
Age:45-49	-.067(.407)	-.074(.379)	-.085(.381)
Feduc	.072***(.023)	.073***(.023)	.073*(.023)
Meduc	-.030(.046)	-.022(.045)	-.023(.045)
Urban	1.17***(.308)	1.07***(.310)	1.02***(.311)
Lnholding	1.05(.744)	.975(.733)	.969(.733)
Hincome	.008(.009)	.008(.009)	.008(.009)
Obs.	925	925	925
R^2	.1408	.1590	.1593
	Dep.var. Mage, sub-sample for *Educ2*		
Avmage	-	.737***(.099)	.726***(.099)
Maithili	-.592***(.1866)	-	-.510***(.185)
Age:30-34	.166(.149)	.155(.146)	.140(.147)
Age:35-39	.062(.148)	.056(.147)	.036(.146)
Age:40-44	.159(.175)	.140(.174)	.116(.174)
Age:45-49	.203(.225)	.181(.223)	.160(.222)
Feduc	.081(.020)	.084(.019)	.084(.019)
Meduc	-.021(.044)	-.015(.044)	-.019(.044)
Urban	1.45***(.271)	1.29(.274)	1.28****(.273)
Lnholding	1.02(.713)	.927(.706)	.919(.704)
Hincome	.009(.010)	.008(.009)	.008(.009)
Obs.	2636	2636	2636
R^2	.1404	.1553	.1569

Notes: Robust standard errors in parentheses. *** significant at 1%, ** significant at 5%, * significant at 10% level. IV1: Maithili. IV2: Avmage.
-Age:25-29 as base category.

Table 2.19: Second-stage regression results: Adult marriage (Mage $\geq$ 16)

Variable	**OLS**	**IV1**	**IV2**	**Two IV**
		Dep.var. *Educ1*		
Mage	.272***(.034)	.253(1.15)	.438**(.223)	.434*(.219)
Age:30-34	-.464*(.261)	-.458(.487)	-.522*(.272)	-.521*(.272)
Age:35-39	-.426*(.249)	-.426*(.246)	-.421*(.251)	-.421*(.250)
Age:40-44	-.601*(.316)	-.604*(.357)	-.573*(.324)	-.574(.323)
Age:45-49	-1.38***(.451)	-1.38***(.450)	-1.37***(.439)	-1.37***(.439)
Feduc	.204***(.021)	.206**(.086)	.193***(.026)	.193***(.026)
Meduc	.066**(.031)	.065(.045)	.071***(.034)	.071***(.034)
Urban	1.84***(.293)	1.86(1.39)	1.64***(.388)	1.65***(.386)
Lnholding	.580(.684)	-.572(1.42)	-.764(.615)	-.760(.616)
Hincome	.001(.003)	.001(.011)	-.000(.002)	-.000(.002)
IVs F-statistic	-	[0.56]	[18.62]	[9.49]
Sargan test p-value				{.8674}
R^2	.3417	.3514	.3224	.3233
		Dep.var. *Educ2*		
Mage	.184***(.025)	1.58***(.556)	.398***(.151)	.514***(.147)
Age:30-34	-.953***(.194)	-1.21***(.297)	-.992***(.197)	-1.01***(.199)
Age:35-39	-1.36***(.180)	-1.48***(.269)	-1.38***(.180)	-1.39***(.183)
Age:40-44	-1.88***(.189)	-2.14***(.328)	-1.92***(.195)	-1.94***(.200)
Age:45-49	-2.11***(.213)	-2.43(.408)	-2.16***(.225)	-2.19***(.233)
Feduc	.434***(.025)	.319***(.055)	.416***(.028)	.407***(.028)
Meduc	.145***(.046)	.169***(.074)	.148(.046)	.150***(.047)
Urban	2.84***(.389)	.788(.937)	2.52***(.385)	2.35***(.387)
Lnholding	.983(.839)	-.471(1.01)	.765(.771)	.643(.737)
Hincome	.007(.009)	-.005(.009)	.005(.008)	.004(.007)
IVs F-statistic	-	[10.07]	[54.49]	[30.08]
Sargan test p-value				{.0035}
R^2	.4557		.4350	.4063

Notes: Robust standard errors in parentheses. *** significant at 1%, ** significant at 5%, * significant at 10% level. IV1: Maithili. IV2: Avmage.
-Age:25-29 as base category.

Table 2.20: First-stage regression results : Four regions

Variable	**IV1**	**IV2**	**Two IV**
	Dep.var. Mage, sub-sample for *Educ1*		
Avmage	-	.665***(.186)	.660***(.185)
Maithili	-1.55***(.518)	-	1.53***(.515)
Age:30-34	-.065(.295)	.024(.293)	-.063(.294)
Age:35-39	-.281(.289)	-.233(.289)	-.272(.288)
Age:40-44	-.107(.359)	-.060(.358)	-.086(.356)
Age:45-49	-.708(.413)	-.657*(.412)	-.728(.411)
Feduc	.093***(.027)	.089***(.027)	.094***(.026)
Meduc	-.033(.051)	-.019(.051)	-.023(.051)
Urban	1.15***(.324)	1.05***(.325)	1.01***(.324)
Lnholding	1.23*(.660)	1.18*(.659)	1.14*(.657)
Hincome	.011*(.007)	.011*(.007)	.010*(.007)
Obs.	960	960	960
R^2	.1896	.1928	.2005
	Dep.var. Mage, sub-sample for *Educ2*		
Avmage	-	.921***(.127)	.951***(.127)
Maithili	-.747***(.211)	-	-.838***(.209)
Age:30-34	-.087(.188)	-.077(.186)	-.104(.186)
Age:35-39	-.327*(.176)	-.318*(.174)	-.344*(.174)
Age:40-44	-.467**(.188)	-.452**(.187)	-.489**(.187)
Age:45-49	-.916***(.206)	-.912***(.205)	-.927***(.204)
Feduc	.118***(.022)	.119***(.022)	.120***(.022)
Meduc	-.006(.048)	.006(.048)	.001(.047)
Urban	1.62***(.250)	1.45***(.250)	1.41***(.250)
Lnholding	1.28**(.649)	1.16**(.645)	1.14*(.644)
Hincome	.013*(.007)	.012*(.007)	.012*(.007)
Obs.	3063	3063	3063
R^2	.2067	.2170	.2211

Notes: Robust standard errors in parentheses. *** significant at 1%, ** significant at 5%, * significant at 10% level. IV1: Maithili. IV2: Avmage.
-Age:25-29 as base category.

Table 2.21: Second-stage regression results: Four regions

Variable	**OLS**	**IV1**	**IV2**	**Two IV**
		Dep.var. *Educ1*		
Mage	.316***(.030)	.422*(.287)	.503**(.245)	.469**(.187)
Age:30-34	-.472*(.253)	-.474*(.253)	-.476*(.257)	-.475*(.255)
Age:35-39	-.519**(.249)	-.493*(.258)	-.473*(.260)	-.481*(.255)
Age:40-44	-.522*(.310)	-.513(.310)	-.507*(.314)	-.510*(.312)
Age:45-49	-1.40***(.407)	-1.34***(.400)	-1.28***(.393)	-1.31***(.378)
Feduc	.194***(.021)	.185***(.034)	.178***(.032)	181***(.028)
Meduc	.073**(.031)	.076*(.045)	.078*(.045)	.077*(.045)
Urban	1.86***(.290)	1.73***(.443)	1.64***(.408)	1.68***(.360)
Lnholding	-.639(.656)	-.779(.678)	-.876(.656)	-.837(.621)
Hincome	.001(.003)	.001(.007)	-.001(.007)	-.001(.006)
IVs F-statistic	-	[9.00]	[12.76]	[10.90]
Sargan test p-value				{.8301}
R^2	.3674	.3578	.3377	.3474
		Dep.var. *Educ2*		
Mage	.194***(.019)	.575**(.272)	.316**(.124)	.370***(.109)
Age:30-34	-.845***(.178)	-.820***(.181)	-.837***(.168)	-.833***(.169)
Age:35-39	-1.38***(.164)	-1.27***(.188)	-1.34***(.161)	-1.33***(.162)
Age:40-44	-1.79***(.170)	-1.62***(.216)	-1.74***(.176)	-1.71***(.176)
Age:45-49	-1.94***(.179)	-1.60***(.316)	-1.83***(.216)	-1.78***(.211)
Feduc	.387***(.025)	.342***(.038)	.372***(.025)	.366***(.024)
Meduc	.195***(.042)	.196***(.046)	.196***(.043)	.196***(.043)
Urban	2.64***(.301)	2.01***(.511)	2.44***(.303)	2.35***(.290)
Lnholding	.900(.761)	.400(.724)	.754(.600)	.681(.601)
Hincome	.010(.010)	.005(.008)	.008(.006)	.008(.006)
IVs F-statistic		[12.49]	[52.49]	[34.36]
Sargan test p-value				{.3740}
R^2	.4545	.3538	.4441	.4329

Notes: Robust standard errors in parentheses. *** significant at 1%, ** significant at 5%, * significant at 10% level. IV1: Maithili. IV2: Avmage.
-Age:25-29 as base category.

Chapter 3

Earnings and Caste: An Evaluation of Caste Wage Differentials in the Nepalese Labour Market (joint with professor Gabriel Montes-Rojas and professor Saqib Jafarey)

3.1 Introduction

A caste system allocates social labour on the basis of a hierarchy of caste classifications and this restricts occupational mobility in two self-perpetuating ways (Banerjee and Knight, 1985). First, caste classification discourages low-caste workers from developing their human capital in line with occupations assigned to the higher castes.

Second, it subjects backward castes to informational and network disadvantages because of their exclusion from certain sectors of employment. Thus, a caste-based division of labour can perpetuate itself through the inter-generational transmission of low educational and occupational status from one generation to the next even once discrimination *per se* is abolished (Borjas, 1994; Darity and Mason, 1998).

This study first examines caste differences in access to large firms and better occupations in Nepal, a country in which, until 1963, an age-old, caste-based social division of labour was imposed by the national legal code *Muluki Ain*. The new *Muluki Ain* of 1963 discarded this caste system. However, caste-based discrimination was itself declared illegal only after the promulgation of the new Constitution of the Kingdom of Nepal in 1990, which made the practice of *untouchability* illegal. Since then, several policies have been implemented to reduce the impact of such discrimination, including positive discrimination and the establishment of a Dalit Commission. The Second Amendment of the Civil Service Act, 1993, reserves 45% of total vacancies in the public sector for backward castes, female, disabled and remote inhabitants. The effects of such policies are not known, partly due to a lack of rigorous research on the subject. This study partly aims to fill this gap.

Secondly, this study examines whether caste differences in labour market outcomes can be attributed to the difference in accessibility discussed above particularly focusing on access to large firms. In doing so, it follows the modern empirical literature and distinguish between *pre-market* and *current market* labour discrimination. The first type of discrimination captures the effects of the propagation mechanisms mentioned above that contribute to the persistence of wage inequality even if active discrimination is no longer practiced by employers. The second type represents active discrimination by employers. The Oaxaca decomposition method (Oaxaca, 1973; Blinder, 1973) is the most commonly used technique for disentangling the

two effects. Empirical studies based on the Oaxaca decomposition have focused on human capital endowments as the sole proxy for pre-market effects (Grimshaw and Rubery, 2002). In addition, Darity and Mason (1998) identifies group differences in access to better paying industries and occupations as major contributors to the persistence of labour market discrimination. Empirical work carried out by Madheswaran and Attewell (2007) and Banerjee and Knight (1985) estimate such effects in the Indian labour market by incorporating occupation in the wage differential decomposition method.

This study goes further in capturing the effects on wage inequality by introducing firm characteristics to supplement educational and occupational differences. In imperfectly competitive markets, firms may remunerate their employees differently, even if they have similar levels of education and work in the same occupation (Vietorisz and Harrison, 1973). In such a situation, employer characteristics such as size, profitability and reputation might matter as much in explaining wage differences as employee characteristics such as education and occupation. This study proxies better paying employers by the size of their firms. The empirical literature provides evidence that larger firms hire higher quality workers (Brown and Medoff, 1989; Schmidt and Zimmermann, 1991; Hettler, 2007).

Accordingly, this study expands the Oaxaca method and estimates three separate decomposition models: one using occupation, another using firm size, and a third in which occupation is interacted with firm size. A relatively larger coefficients for access to job components in the third decomoposition model (interaction between occupation and firm size) implies that differences in access to large firm positively contributes on differences in labour market outcomes across castes. Similarly, a relative reduction in it's magnitudes in the latter period of analysis indicates the effectiveness of government policies of positive discrimination in promoting access

to public sector job by historically backward castes.

Results indicate that caste wage discrimination is indeed present in the Nepalese labour market, with intermediate (*Matwali*) and low (*Pani Nachalne*) castes earning significantly less than the higher (*Tagadahari*) castes. Moreover, lack of access to employment in certain occupations and larger firms is found as a major factor behind the caste wage differential together with years of schooling.

The rest of the study is organized as follows. Section 3.2 reviews previous literature on labour market discrimination. Section 3.3 describes the historical and institutional basis of caste classification in Nepal. Data and estimation strategy are discussed in Section 3.4 while Section 3.5 states the econometric models. The main econometric results are presented in Section 3.6. Section 3.7 concludes.

3.2 Literature review

3.2.1 Labour market discrimination: Basic concepts

Labour market discrimination is defined as the market valuation of workers' personal characteristics, such as age, ethnicity, sex etc., which is not related to productivity (Arrow, 1971). Autor (2003) describes it as the situation of minority workers being treated less favourably than the workers from dominant groups with identical productivity characteristics. Therefore, labour market discrimination can explicitly be defined as a situation in which a person who provides labour market services and is equally productive in a physical and material sense is paid less in a way that is related to caste or ethnicity (Altonji and Blank, 1999).

Discrimination in the labour market reinforces income inequality across groups in societies (Cain, 1984). It can lead to the occupational and industry-type segregation of labour forces (Autor, 2003). Since workers from different social groups are not

paid according to their marginal productivity (MP) it can also be considered as an underlining factor of market failure and thus to produce economic inefficiency in an economy.

Despite these negative impacts, labour market discriminations have been constantly witnessed in virtually all regions in the world irrespective of their levels of socioeconomic development. This has led to the widening interest of researchers to investigate underlying factors of labour market discrimination and its persistence. As a result a bulk of literature on this subject, theoretical as well as empirical, has already been documented.

The classical view of economics highlights that the total output of an economy will be distributed to the factors of production, labour and capital, according to their MP. If this classical view holds then the group differences in wage earnings arise from differences in MP, actual or perceived by employers. Additionally, employers may depict varied tastes over identically productive workers from different groups. Difference in perceived MP as well as varried tastes by employers over the workers from different social groups fabricates the labour market discrimination.

A hypothetical example bellow illustrates the effect of discrimination on hourly wage rate between the workers from dominant and minority ethnic groups employing a simple labour demand and supply framework. Labour supply for the both types of worker in this example is assumed to be fixed.

Figure 3.1 depicts two types of labour supply namely dominant and minority with identical productivity characteristics. In this example, employer's perceived MP for minority workers is less than that for workers from the dominant group. It can also be assumed that employers have distaste over minority workers. As a result, demand for minority workers, Q^*m, is lower than the demand for dominant workers Q^*d although availability of labour (labour supply) is same for both groups.

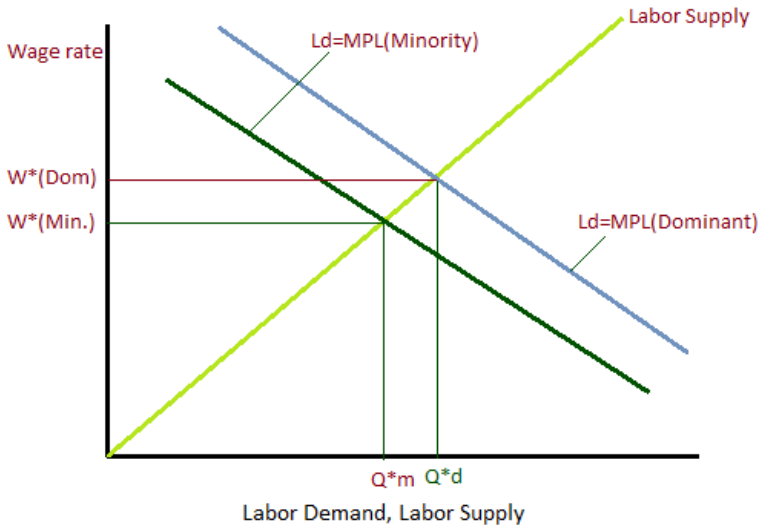

Figure 3.1: A basic model of labour market discrimination, Adapted from Riley (2012)

A relatively less demand for minority workers leads to a lower equilibrium wage rate, $w^*(min)$ for them compared to the equilibrium wage rate associated with the workers from dominant group, $w^*(dom)$. It compels minority workers either to increase their productivity level relative to dominant worker in order to achieve the same rate of wage or accept a lower wage rate than dominant workers for identical levels of productivity.

3.2.2 Theories of discrimination in labour markets

Various approaches have been documented in explaining discrimination in the labour market which can broadly be classified into competitive and collective models. However, the majority of studies in the labour market discrimination are based on competitive model where employers maximize their utility not only from financial profits but may also include the economic value of "taste" (Autor, 2003). Widely used com-

petitive models of discrimination are taste discrimination (Becker 1957, 1971) and statistical discrimination (Phelps, 1972; Arrow, 1971).

3.2.2.1 Taste discrimination

The theory of taste discrimination (Becker 1957, 1971) incorporates employers' prejudicial behaviour in hiring and wage setting process for their employees. Employers are assumed to possess varied tastes over different groups of employee in this framework. In other words, employers' distaste against particular group of workers enters into their utility function as disutility. This model can be explained by a simple formulation of the behaviour of profit maximizing firm.

Let a firm produces its output utilising single factor of production, labour, which comprises of workers from minority (discriminated-against) and dominant ethnic groups. Firm's objective function can be characterized as,

$$U = PF(N_d + N_m) - w(N_d + N_m) - \delta N_m \tag{3.1}$$

where, p is the price level, F is production function, N_d and N_m are number of workers representing respectively from dominant and minority groups. w and δ represent wage rate and discrimination coefficient, respectively. This equation shows that firm's profit depends on total revenue generated from output produced, wage paid to labours (comprising of two types of labours) and the discrimination coefficient associated with minority workers, δN_m.

First order condition (FOC) gives the optimal number of labour from each group that firm will be willing to buy in order to maximize profit

$$\frac{\partial U}{\partial N_d} = pF'(N_d) - w_d = 0 \tag{3.2}$$

and

$$\frac{\partial U}{\partial N_m} = pF'(N_m) - w_m - \delta = 0 \tag{3.3}$$

Equations 3.2 and 3.3 show that prejudiced employers will hire worker from minority groups only if $w_d \geq w_m + \delta$ which indicates labour market discrimination against the minority group of workers. However, under the assumption of competitive labour market and constant return to scale (CRS) non-discriminatory employers will hire workers for minority groups to expand their business. In equilibrium discriminatory employers will pay the price of their distaste since they cannot pass this cost to the customers. This in turn leads prejudiced employers into insolvency with entry of non-discriminatory employers in labour markets or they no longer discriminate in order to survive in the competitive market. Therefore, under a competitive market, taste discrimination is not likely to be persist rather it seems to be a short-run phenomenon.

Additionally, the fraction of prejudiced employers should be sufficiently large in order to produce wage discrimination in the labour market. If labour market comprises of a small number of discriminatory employers, workers from minority groups may choose to work for other non-discriminatory employers and thus discriminatory effect will be competed away. It is likely to create incentive for segregation rather than to produce labour market discrimination. Therefore, as long as the both groups of workers possess a similar average level of productivity they will eventually be paid the same wage. Resulting difference will only be the segregation of workforce, i.e., black workers will work for black employers and white workers for white employers which is not captured in Becker's model of taste discrimination (Darity and Mason, 1998).

3.2.2.2 Statistical discrimination

The majority of literatures in labour economic focus on statistical discrimination (Phelps, 1972; Arrow, 1971) than the theory of "taste discrimination" (Becker, 1957, 1971) in explaining observed wage gaps in the labour market (Autor, 2003). The premises of this approach is that employers possess limited sources of information in predicting the productivity of potential employees that compels them to rely on some easily observable group characteristics such as race, gender, ethnicity and caste in inferring the productivity of prospective workers. Therefore, employers consider group association rather than individual attributes while perceiving employees' productivity in this framework.

The notion of statistical discrimination can be explained by a simple algebraic formulation. Suppose that dominant and minority workers have individual test score s_j (j=d,m). If test score perfectly explains the individual productivity they will eventually be paid according to their MP. However, employers lack information on exact productivity associated with prospective employees and therefore take into account their perceived average productivity, S_j^*, which varies across groups. Expected productivity and thus the wage that employers will be offering to the both types of workers now can be expressed as,

$$W_d = \alpha S_d + (1 - \alpha) S_d^* \tag{3.4}$$

$$W_m = \alpha S_m + (1 - \alpha) S_m^* \tag{3.5}$$

where, α represents the correlation between actual productivity and test score. These equations show that not only the perceived mean productivity associated with different groups of workers but also variance of true productivity affects the wage rate. More importantly, since the productivity is perceived by employers but

not real they will know employees' actual productivity in due course of action and thus will eventually update their belief. This, once again, shows that labour market discrimination does not persist for the long-run.

3.2.2.3 Customer discrimination

If the majority of customers show prejudiced behaviour over minority groups they will be biased in buying goods and services produced by minority workers. In other words customers perceived utility adjusted price (adjusted for the discriminatory coefficients) for the goods and services sold by minority workers will be less than that of goods and services sold by the workers from dominant groups. Employers thus have to charge fewer prices for the commodity produced by minority workers in order to compensate customers disutility. This in turn will be transmitted into the wage rate of minority workers since employers will not be willing to borne out such cost from their own profit. This shows that labour market discrimination can also be aroused from the prejudiced behaviour by customers. However, there will be a possibility that minority workers might chose to work for employers where customers direct contact is not essential in order to avoid the impact of customers distaste on their wage. It implies that customer discrimination may lead to the workforce segregation but not necessarily to the labour market discrimination.

3.2.3 Empirical evidence of labour market discrimination

A range of empirical studies have shown evidences that gender or ethnic discrimination exists in labour markets. A comparative study of 7 Latin American countries (Bolivia, Brazil, Chile, Ecuador, Guatemala, Paraguay and Peru) shows significant labour market discrimination based on ethnicity that is mainly coming from differences in educational attainment (Atal et al., 2009). In Bolivia indigenous Spanish

speaking workers were found to earn less relative to their non-indigenous counterparts of which more than 50 percent attributable to the pre-market discrimination (Atal et al., 2009). A study of racial and gender wage disparities in urban Sao Paulo, Brazil, found that ethnic minorities (Afro-Brazilian) and women receive lower salaries compared to their white male peers (Lovell, 2006). Differences in human capital endowments explained only a small part of the wage gaps. This study identified an increasing trend in the unexplained portion of wage gaps over time. Lerrea and Torres (2006) examined socioeconomic and ethnic determinants of earnings in Ecuador. They conclude that indigenous workers receive 55 percent less wages than non-indigenous workers. In line with others findings, their study also finds that most of the wage differential to be explained by differences in educational levels. Differences in levels of schooling as well as disparities on returns to education associated with different ethnicities are found in Guatemala (Patrinos, 1997). Queche receives low return to education compared to Mam counterparts in the Guatemalan labour market. A study carried out in South African labour market pointed out $30-50$ percent racial wage gaps (between black and white workers) where roughly $40-50$ percent of which is due to the unequal endowments and possibly clustering of black workers in particular occupation and industries (Szelewicki and Joanna, 2009).

3.2.4 Persistence of labour market discrimination

As stated before, theories of taste discrimination and statistical discrimination are two major theoretical explanations for labour market discrimination. The former explains that employers directly hold preferences about the ethnic background of their employees (Becker, 1957, 1971) while the latter highlights on employers incomplete information about workers' productivity and statistical priors about how productivity varies with ethnicity (Aarrow, 1971; Akerlof, 1984; Phelps, 1972). Both

of these theories are based on a neoclassical model that implies that competition will lead to the elimination of race or gender discrimination in the long run. Under perfect competition and different degrees of prejudicial tastes, employers with the least taste for discrimination will hire members of a minority group but pay them a lower wage than equally qualified non-minorities (Chase, 2000). By paying less than the equilibrium wage, those employers earn profits in the short run. In the long run, however, taste discrimination disappears with insolvency of prejudiced firms and new entries of less prejudiced competitors into the market. Similarly, if group differences in ability are perceived to exist by employers but are not real, as the theory of statistical discrimination assumes, employers will update their beliefs over time (Darity and Mason, 1998). However, evidences from empirical works carried out in various regions and over different points of time have shown that labour market discrimination persists. It indicates that traditional approaches, taste discrimination as well as statistical discrimination, fail to explain the persistence nature of labour market discrimination and necessitates alternative approaches to analyse this phenomenon.

3.2.4.1 Labour market imperfection

Wage rate in the labour market is not always set according to workers productivity as the classical theory of marginal productivity explains. In fact, it can be influenced by various non-market mechanisms. For instance, trade union can influence employer to fix their wage higher than the equilibrium level. Similarly, monopsony in the labour market can discriminate certain groups of workers paying them less than their MP. In monopsony, single buyer can offer wage less than MP to certain groups of worker. This situation is likely to prevail when a large firm dominates the local labour market (Chase, 2000). Additionally, institutional factors which restrict occupational and

geographical mobility of workers might reinforce monopsony. Caste-based social division of labour or an official imposition on geographical restriction of workers' mobility are some of its examples. A case study of Latvian labour market during post- communist transition shows that Russians suffer labour market discrimination between 5.5% to 7.3% relative to Latvians assuming the former to have lower labour supply elasticity as the effects of monopsony (Chase, 2000).

An additional factor to perpetuate labour market discrimination is imperfect competition across employers (Botwinick, 1993; Darity and Williams, 1985). These authors argue that, in the absence of current discrimination, perfect competition would drive employers to offer workers of a given educational level and occupation the same wages, under imperfect competition, more profitable employers would be able to offer higher wages to employees, despite similarities in education, occupation and industry-type. This approach emphasizes how inequality across different groups of workers in access to jobs with employers who pay higher wages (Darity and Mason, 1998) can persist over time as a result of past discrimination and suggests that the wage structure across social groups should be studied as a joint function of individual and employer characteristics.

3.2.4.2 Induced deficiency in human capital endowment

Labour market discrimination can have a prolonging effect on the labour market outcome of discriminated-against groups via pre-market effect. In other words, workers from discriminated groups may inherit low level of human capital endowment relative to those who enjoyed labour market premia in the past. Therefore, some authors have looked into the *persistence* of discrimination caused by "pre-" or "extra-market" processes, which reduce the earning power of previously discriminated-against groups (Madden, 1975). According to this view, the main

type of 'pre-market' effect is the inter-generational transmission of low levels of education. The reason is that when a generation of workers grows up suffering directly from discrimination, not only is their own education undermined, but even if the discrimination ends during their lifetime, they have both less incentive (via higher discounting of the returns to education) and fewer means (via lower income) to invest in the education of their own children (Agee and Crocker, 1996). Carneiro et al. (2005) argue that future expectation of labour market outcome formed by past experience induces parents as well as children to invest less in skills development activities. Moreover, because less educated parents are less capable of helping their offspring with schoolwork than are their more educated, higher income counterparts, children from disadvantaged groups are likely to fare worse in schooling outcome than their privileged counter-parts (Bond, 1981). Ermisch and Francesconi (2001) also demonstrate that parental educational background is a significant determinant of a child's educational achievement.

Other authors have emphasized the effect of that low levels of inter-generational transmission of educational attainment have on limiting occupational mobility across generations (Checchi, 1997). This can become an additional factor along with low educational endowments *per se* in reinforcing persistence of labour market discrimination. Madheswaran and Attewell (2007) and Banerjee and Knight (1985) both find evidence in favour of this hypothesis. Therefore, induced or inherent deficiency within the groups that experience the inferior economic outcome can be viewed as a main cause of perpetuating labour market discrimination (Darity and Mason, 1998).

3.2.5 Labour market discrimination and caste system

While considerable attention has been paid to labour market discrimination based on race and gender, less attention has been paid to the issue of caste even though

caste-based discrimination might be more powerful and persistent than racial discrimination. Racism emerged in countries that were either colonized or participated in the slave trade during the colonial era, while caste-based societies have existed for centuries before colonialism (Deshpande, 2011). Moreover while, apart from the master-slave division of slavery, the colonial powers did not impose strict occupational restrictions on the subject population, caste-based stratification was inherently associated with an occupational division of labour.

Akerlof (1984) describes a caste-segregated society as being more self-perpetuating than one in which current discrimination is of the 'tastes' or 'statistical' variety. The reason for this is a self-enforcing mechanism inherent in the caste system, whereby third parties punish any employer-employee pair who deviates from the caste division, by casting out both from normal society. He shows that caste equilibrium exists in which no single member of either the dominant or disadvantaged group is willing to break away from the caste division of labour, despite the presence of potential bilateral gains between employers and employees from doing so. In practice, members of disadvantaged castes also adhere to the caste structure out of self-fulfilling sense of fatalism (Kuran, 1987). Thus, even after countries such as India and Nepal legally outlawed caste discrimination, the tendency for caste discrimination to persist through decentralized behaviour continued.

3.2.6 Empirical studies on caste discrimination in labour markets

Das and Dutta (2007) estimated the caste wage differential in both regular and casual jobs in the Indian labour market. Their results showed that a substantial differential existed between Scheduled Castes (SC) and General Castes (GC) in regular jobs, but not in casual ones, with almost two third of the differential in regular

jobs being attributable to endowment effects. However, while their endowment variable included both educational and occupational proxies, amongst others, it did not explicitly differentiate between different types of endowment; therefore, their study cannot explicitly report the impact of specific types of endowment differences on the caste wage differential.

In a similar study of the wage differential between Scheduled and non-Scheduled caste migrants in Delhi, Banerjee and Knight (1985) found that low-caste workers were more likely to be engaged in traditional low-paid jobs. By extending the conventional decomposition methodology to include occupational access as part of a worker's pre-market endowment, they found that a significant part of the caste wage differential was attributable to difference in access to better paid occupations. In a study of regular salaried jobs in India, Madheswaran and Attewell (2007) found that endowment differences were more significant than current market discrimination in explaining the caste wage differential, and further, that the most important type of difference in endowments was the difference in occupation across castes.

These studies highlight the fact that pre-market effects of discrimination are more important than current market discrimination in explaining the persistence of the caste wage differential in India, with the difference in access to better paid occupations as a major factor in promoting pre-market differences.

3.2.7 Contribution of this study

Studies on caste wage differential have taken into account the modern empirical literature on the subject, which distinguishes between 'current market' and 'pre-market' labour discrimination. The later type of effect is known in the literature to contribute to the persistence of wage inequality even if active discrimination is no longer in current practice by employers. The Oaxaca (1973) decomposition method

is the most commonly used technique for disentangling 'pre-market' from 'current market' discrimination. Conventional Oaxaca decomposition methodology focuses on the human capital endowment (Grimshaw and Rubery, 2002). Therefore, the majority of empirical works consider productivity characteristics (education, experience) as the sole proxy for pre-market effects. Additionally, Darity and Mason (1998) identifies 'group' differences in access to better paying jobs within industries and occupations as major contributive factors of persistence in labour market discrimination. Empirical work carried out by Madheswaran and Attewell (2007) and Banerjee and Knight (1985) estimate such effects in the Indian labour market by incorporating occupation in the wage differential decomposition method.

This study questions the adequacy of using only education and occupation as proxies for pre-market effects in countries that have imperfectly competitive markets for both labour and goods and services. The reason is that, in imperfectly competitive markets, there might exist a hierarchy of employers that pay differently to workers of the same educational level and occupational classification. For instance, in Nepal, where the public sector is a significant employer of white-collar workers (Moore, 2006) jobs are categorized by grade, with each grade associated with its own point in the overall pay scale. Engineers, doctors, administrators or accountants falling into a given grade all receive similar wages. However, the private sector may display segmentation within the same market, so that different firms might display different levels of both profitability and ability to pay higher wages (Bluestone, 1974; Vietorisz and Harrison, 1973).

This study hypothesize that workers from marginalized castes inherit limited occupational choices as well as unequal access to better firms, proxied by firm size. In such a situation, workers' characteristics such as education and occupation might matter less than employer characteristics in explaining observed differences in labour

market outcomes. Acknowledging this, better paid jobs proxied by firm size has been employed in this study to analyse the source of caste wage differential. The empirical literatures have provided evidences that larger firms hire higher quality workers (Borwn and Medoff, 1989). Similarly, Schmidt and Zimmermann (1991) and Hettler (2007) have shown a positive relationship between firm size and the wage rate.

This hypothesis is tested by analysing the caste differences in the probability of obtaining better occupations and being employed in large firms. Accordingly, this study first evaluates caste differences in the probability of obtaining better occupations and access to large firms. Secondly, it expands the Oaxaca decomposition method to use occupation, firm size and their interaction as indicators of caste-specific 'endowments' rather than as control variables. Therefore, three separate decomposition models are estimated in investigating the sources of caste wage differentials in the labour market. A larger magnitude of access to job component in interaction model compared to other two models implies that cast wage differentials are coming from the caste differences in access to larger firms. Therefore, main methodological contribution of this work is to use an expanded set of proxies for detecting the presence of such pre-market effects.

Moreover, an empirical study on Nepalese labour market has not been carried out as yet, even along the lines of the studies on India discussed above. This is an important gap which needs to be filled in light of the fact that Nepal is the only country to have had a long history of a *legally* imposed caste based division of labour. Caste-based social division of labour was imposed in Nepal by the national legal code *Muluki Ain* until 1963. The new *Muluki Ain* of 1963 discarded this caste system. However, caste based discrimination was itself declared illegal only after the promulgation of a new Constitution of the Kingdom of Nepal in 1990

(this made the practice of *untouchability* illegal). Since then, several policies have been implemented to reduce the impact of such discrimination, including positive discrimination and the establishment of a Dalit Commission. The Second Amendment of the Civil Service Act, 1993, reserves 45% of total vacancies in the public sector for backward castes, female, disabled and remote inhabitants.[1] The effects of such policies are not known, partly due to the lack of rigorous research on the subject. Thus, an additional scope of this study is to offer a statistical analysis on the effectiveness of these policy interventions. A relative decrease in wage differential due to the group difference in access to better paid jobs, larger firm, will indicate the positive contribution of this policy.

Econometric results indicate that wage discrimination is indeed present in the Nepalese labour market, with intermediate and low castes earning significantly less than the higher castes (section 3.3 offers a precise classification of the different castes used in this analysis). Moreover, lack of access by intermediate and low castes to employment in larger firms is found as an important factor behind the caste wage differential along with productivity characteristics, years of schooling.

3.3 Caste system in Nepal: An overview

Nepal, along with other countries, had a caste-based social division of labour in the past. Historically, caste classification in the Indian sub-continent was based on the Varna system of Hindu philosophy and the Aryan division of labour. These comprised four categories, namely Brahman, Ksyatriyas (Chhetri), Vaisyas and Shudras. Together these encompassed a social division of labour as priests and teachers, warier and royalty, merchants and money lenders, and artisans, service providers and other menial workers, respectively (Deshpande, 2011). Brahman, being the superior

[1]This ammendment was made on August, 2007.

caste, enjoyed the best status in the Nepalese society followed by Chhetri. While Vaisyas did not seem as privileged as Brahmans or Chhetri, they enjoyed a relatively higher social status than Shudras on the caste based social hierarchy. Shudras were the lowest caste, considered as untouchable by their superiors.

As a predominantly Hindu country with a significant Buddhist minority, Nepal's adopted the Hindu caste system came with some local variation, implemented in the form of a legal code called *Muluki Ain*. This code classified all Nepalese into different categories irrespective of their religious backgrounds, but based on their relative ritual purity (Bennett et al., 2008). The official classification under *Muluki Ain* consisted of three categories, namely *Tagadhari* (literally "twice-born"), *Matwali* (literally "liquor drinking") and *Pani Nachalne* (literally "impure") (Cox, 1988). *Tagadahari* included upper-caste Hindus such as the Brahmans of the traditional Hindu caste system. *Matwali*, on the other hand, consisted mainly of Buddhists and indigenous ethnic groups who practiced Animism and Shamanism, and were considered an intermediate caste. The *Pani Nachalne* were the lowest caste and included not just traditional Hindu untouchables such as Kami, Sarki, etc. but also Muslims and Mlechha (literally 'foreigners'), who in turn included Christians.

This is where the intersection of caste and ethnicity entered into the social hierarchy of Nepal. Hofer (1979) and Gurung (2002) describe a hierarchy of ethnic groups and their respective association with the legal caste categories. This divides all ethnic groups into two broader categories of Pure and Impure caste hierarchies consisting of three and two subcategories, respectively (see Table 3.1).

Figure 3.2 portrays Nepalese caste hierarchy. While ethnic groups belonging to the *Tagadhari* and *Matwali* castes, fell under Pure (or Water Ac-

Figure 3.2: Caste hierarchy

ceptable, i.e. sharing water with them was acceptable), the *Pani Nachalne* were Impure (or Water Unacceptable). Within these there were sub-categories: while the Pure *Matwali* were divided into Enslavable and Non-enslavable ethnic groups, the Impure *Pani Nachalne* were further divided into Untouchable and Touchable, depending on whether or not they belonged to Hindu religious groups.

In line with these classifications, this analysis aggregates caste-ethnic identity into three broad categories, namely *Tagadhari*, *Matwali* and *Pani Nachalne*. However, lack of observations on the Enslavable *Matwali* and Touchable *Pani Nachalne* groups has prevented in constructing a finer division of the social hierarchy. These are referred as castes although from a strict point of view they correspond to caste and ethnicity.

This study estimates the impact of labour-market endowments and job characteristics, human capital (education), occupation, and firm size, on caste wage differentials by applying Blinder (1973) and Oaxaca (1973) decomposition technique and the model extended by Banerjee and Knight (1985). The decomposition methodology is further expanded in this study to incorporate occupation, firm size and their interaction into the model.

3.4 Data and estimation strategy

3.4.1 Data and descriptive statistics

This study employs two waves of the National Living Standard Survey (NLSS) of Nepal for 2003/2004 and 2010/2011 carried out by the Central Bureau of Statistics

of Nepal with the combined support of the World Bank and the UK Department for International Development (these surveys will be referred below as 2003 and 2010, respectively.) The surveys follow the World Bank's Living Standard Measurement Survey and apply a two-stage sampling scheme. 73 out of the 75 administrative districts of Nepal are covered. A total of 5240 households in 2003 and 5998 households in 2010 were interviewed, and information recorded about 28110 and 28670 individuals in each of the respective years. The data include information on wage employment, self-employment, sector of employment, industry type, mode of payment, labor market attachment and educational attainment at the individual level. Since information on experience is not reported, it is proxied by age minus years of schooling minus six, which is the average age to start school in the Nepalese education system. For simplicity, it is assumed that every person joined the labor market immediately after completing their schooling. An individual is defined as employed if he/she worked at least one hour during the seven days prior to the interview. See the Appendix 2 for the details of these classification plus definitions of all variables.

The analysis includes 785 in 2003 and 834 in 2010 male wage workers aged 19-59 years old from the non-agricultural sector.[2] Descriptive statistics are presented in Tables 3.3 and 3.4.

The *Tagadhari* group represents the dominant share of employees in both periods, accounting for 70.7% of the total employment in 2003 and 71.3% in 2010. The *Matwali* accounted for 19.2% and 21.4%, and *Pani Nachalne* 9.9% and 7.3% in each survey year, respectively.

Tables 3.3 and 3.4 show an average log hourly wage rate of 3.34 and 3.89 NPR respectively. The USD equivalent would be .38 and .68, respectively.[3] The *Matwali*

[2]The NLSS has separate questions for agriculture and non-agriculture wage employment. This study only considers respondents in the non-agriculture employment. However, agriculture can also be a selected as an industry in the non-agriculture wage employment questionnaire.

[3]Average exchange rates between NPR and USD were 73.99 and 71.80 in 2003 and 2010, re-

and *Pani Nachalne* workers earn on average wage 30% and 49% less than *Tagadhari* workers, respectively, in 2003. By 2010, the wage gap between the *Tagadhari* and *Pani Nachalne* remains identical whereas it has been descreased to 20% in case of the wage gap between the *Tagadhari* and *Matwali* workers.

Average years of education, defined as the highest level of completed years of schooling were 7.78 in 2003 and 9.88 in 2010. The education gap between *Tagadhari* and *Matwali* was 2.29 years in 2003 and by 2010 it had decreased slightly to 2.10 years. However, the educational gap between *Tagadhari* and the lowest caste *Pani Nachalne* increased over this period, from 3.03 years in 2003 to 4.45 years in 2010.

The NLSS survey contains a question about the size of the firm where the wage worker works. As described in the Appendix 2 it contains three categories: 1 employee, 2-10 employees, and more than 10 employees. This study uses the *ad-hoc* classification of small, medium and large firms, respectively. This variable has a high proportion of missing observations, i.e. non-respondents, which resulted in a particular distribution of workers across occupations. In the robustness section below this study considers the imputation of firm size to certain occupations.

Occupations are aggregated into seven broad groups based on Nepal's *National Classification of Occupations*: professional, clerical, service, skilled, sales, agriworker and unskilled. The professional category includes the categories of doctor, engineer, manager, religious and clerical comprises of categories such as clerk, typist, book keeper, etc. Those not included in any of the six occupations are classified as unskilled workers which in turn includes loaders, unskilled construction workers and laborers. Similarly, eight categories of industry are constructed based on the *Standard Industrial Classification* (SIC) reported in the survey.

In 2003, the occupational ranking is as follows: professionals is the largest cate-

spectively. Source: Nepal Rastra Bank.

gory accounting for 38.6% of workers, unskilled is second largest with 18.4% followed by skilled workers at 17.9%. By 2010 the rankings are 28.2% for skilled, 23.9% for professional and 19.1% for clerical. The professional and clerical occupations, which collectively correspond to white collar jobs, have a higher proportion of *Tagadhari* workers, while the lower castes *Matwali* and *Pani Nachalne* workers are more engaged in unskilled and skilled occupations. In order to highlight the role of firm size, Table 3.5 report average wages in 2003 and 2010 by occupation in the three firm size categories considered in this analysis. In all cases, larger firms pay higher wages than smaller ones.

In terms of the workers' industry, the majority of workers are in the service, manufacturing and other industry classification.[4] There are no significant differences between the *Tagadhari* and *Matwali* workers with respect to their association to industries. The *Pani Nachalne* workers are more likely to work in the manufacturing industry.[5] Information is not available to distinguish between public and private sector employees.

In summary, the descriptive statistics indicate that caste-based disparities in key labor market outcomes continue to play an important role in Nepal. The intermediate *Matwali* group have managed to slightly close the gap with the dominant *Tagadhari* group, while the lowest caste *Pani Nachalne* appears to have fallen further behind. However, the descriptive statistics alone cannot tell the key drivers of these disparities.

[4]Note that there is a significant change in the industry classification between 2003 and 2010 regarding the Other category, which represents industry not responded or responded as other.

[5]Workers from this caste do not have representation in the FRE industry in both periods and FRE, mining and agricultural industries in 2010.

3.4.2 Methodology

Probit and multinomial regression models are widely used methodologies in empirical works in order to estimate the probability of an observation with certain characteristics will obtain specific outcome from given alternatives. The former model is used when there is only two possible outcomes and thus is also known as an ordinal or binary response model while the latter is used when there are more than two categorical variable as possible outcomes. This study uses both models for the first part of study in analysing caste differences in access to large firms and better occupations.

Following the analysis of caste differences in access to large firms and better occupation this study evaluates whether such differences are contributing factors in producing caste differentials in labour market outcome. Mincerian wage equation is the one of the most widely used methodology in empirical economics which defines workers' labour market outcomes as the function of their productivity characteristics. In this case, estimating a wage equation using productivity characteristics along with caste binary variables as explanatory variables can be used in analysing wage differentials across castes. However, this single equation technique can produce biased results since it assumes a similar wage structure for all castes (Madheswaran and Attewell , 2007). The second approach is to use a decomposition technique developed by Blinder (1973) and Oaxaca (1973) which decomposes wage gaps between groups of workers into the endowment and coefficient effects. The former is known as the *pre-market* effect while the latter is considered as *current market* effect in the labour market literature.

However, some authors have argued that occupational distribution plays an important role in explaining a worker's wage. Individuals occupational choice may depends on expected life time earnings, cost involved in acquiring skills suitable for

a particular occupation (Boskin, 1974), race, sex, education and experience (Schmidt and Strauss, 1975), employers willingness to hire or individual desires to work on that occupation (Brown et al. 1980). Accordingly several empirical works have incorporated occupational differences across groups of worker and have treated occupational attainment as an endowment while estimating labour market discrimination. This approach is also known as the expanded decomposition technique which allows distinguishing between wage and job discrimination. The conventional Oaxaca (1973) decomposition disentangles gross wage differential into wage explained and wage unexplained (discrimination) components whereas the expanded decomposition technique decomposes it into four components: wage explained, wage unexplained, job explained and job unexplained. The latter two components correspond to the effects of group differences in access to better occupations on gross wage differentials.

This study intends to contribute to the literature by incorporating firm size distribution along with productivity and occupational characteristics in estimating the source of caste wage differential .[6] Therefore, this study estimates augmented Mincerian wage equations separately for different castes followed by an expanded decomposition technique to estimate different sources of wage differentials.

3.4.3 Variables specification

In the first part, a binary variable for firm size and a categorical variable to represent seven types of occupation are used as dependent variables in order to analyse caste differences in access to larger firms and occupation, respectively. The logarithm of hourly wage is used as the dependent variable in the second part of the study.

[6]However, the explained and unexplained components of job effects are not estimated since main focus of this study is to evaluate the impact of access to large firm on caste wage differentials rather than to focus on caste wage discrimination. Therefore, explained component of job characteristics (occupation and firm size) includes both the explained and unexplained components in expanded decomposition technique discussed above.

Worker's individual characteristics that might influence wage such as education, experience and experience square are included as predictors. Since marital prospect can influence an individual's attitudes toward work (a common belief is that married men hold stronger attitude toward work than the single men) a married dummy variable has also been included as an explanatory variable. Empirical studies have shown that the rate of wage can significantly vary across regions. Therefore, regional dummies are used as additional control variables.[7] Other predictors used in this analysis are rural, occupation, firm-size and industry-type dummy variables (see Appendix 2 for details of variable specification).

Commonly used control variables in wage equations such as age and gender dummy variables are not used in this analysis. Age variable is dropped because of multicollinearity since experience in this study has been proxied by age minus years of schooling minus six (which is an average years to start school in Nepalese education system). Gender dummy variable is not applicable since this study uses a sample of only working-age male workers.[8]

Although it has been argued that caste system restricts occupational mobility it may not completely explain an individual worker's occupational association. In this situation occupation variables may be endogenous. However, this is not considered as a crucial limitation at least for this study since its main focus is to estimate the source of caste wage differential arguing that differences in access to larger firm as its contributing factor. Additionally, several sudies have used occupational categories as predictors while estimating the wage differentials across groups of workers (see for

[7]As very few respondents reported their work place in the first survey and the second survey does not have such information a regional dummy variable (individual's household belonging to particular administrative region) is used as proxy for work place.

[8]Female are excluded in this analysis for two reasons. First, female labour force participation is negligible in developing countries like Nepal. Secondly, they are mostly in informal employment (domestic workers) and thus may not be relevant in assessing the firm size effect on wage differential as this study is focused on.

example Banerjee and Knight, 1985; Hinks and Wastson, 2001 and Madheswaran and Attewell, 2007). Thus it compares the impact of access to occupation, large firm and larger firm with better occupation on caste wage differentials.

3.5 Empirical model

This study first evaluates the caste differences in access to large firms and occupations. It then estimates whether caste wage differentials can be attributed to these differences in accessibility particularly focusing on the access to large firms. A probit model is estimated to look into the caste differences in access to large firms whereas a multinomial analysis has been carried out to estimate the accessibility in better occupation by castes. Oaxaca decomposition model is employed to estimate the latter effects. Considering caste categories $j= t, m, p$ (*Tagadhari*= t, *Matwali*= m and *Pani Nachalne* = p), these models can be specified as,

$$P(y = 1|X) = \alpha_{ij} + X_{ij} + \lambda C_{ij} + u_{ij} \tag{3.6}$$

where, y is a binary variable carrying value one for a large firm and zero otherwise. X_{ij} comprises a set of explanatory variables including education, experience, experience square, marital status and rural dummies, land holding, regional, occupation and industry- type dummies associated with individual i. C_{ij} represents individual i's association with caste j and thus the sign and significance of λ indicates the caste difference in the probability of access in large firms. u_{ij} represents unobserved error terms.

$$P(y = s|X) = \alpha_{ij} + X_{ij} + \lambda C_{ij} + u_{ij} \tag{3.7}$$

where, y is a categorical variable to represent seven different occupations (s=1,2.....7). In this equation X_{ij} includes only productivity characteristics (education and experiance), rural dummies and land holding. Regional, firm size and industry-type dummy variables are not used in order to achieve conversence to estimate the multinomial model.

In the second part it proceeds to estimate the source of caste wage differentials. An expanded Mincerian log wage equation can be specified for each caste as,

$$w_{ij} = \beta_j E_{ij} + \delta_j S_{ij} + \gamma_j X_{ij} + \epsilon_{ij}, \tag{3.8}$$

where w_{ij} is the log hourly wage of individual i of caste j, E_{ij} represents years of schooling completed, S_{ij} is a set of variables containing job characteristics such as occupation and/or firm size (see below), X_{ij} is a set of covariates comprising of a constant, experience, experience square land holding, marital status, regional, rural and industry dummies, and ϵ is the unobserved component in the wage equation.

The gross logarithmic caste wage differentials in observable variables can be viewed as the sum of caste differences in educational effects, access to job effects and the effects of other control variables including constant terms and can be specified as,

$$\overline{w}_t - \overline{w}_m = (\beta_t \overline{E}_t - \beta_m \overline{E}_m) + (\delta_t \overline{S}_t - \delta_m \overline{S}_m) + (\gamma_t \overline{X}_t - \gamma_m \overline{X}_m), \tag{3.9}$$

$$\overline{w}_t - \overline{w}_p = (\beta_t \overline{E}_t - \beta_p \overline{E}_p) + (\delta_t \overline{S}_t - \delta_p \overline{S}_p) + (\gamma_t \overline{X}_t - \gamma_p \overline{X}_p), \tag{3.10}$$

where $\bar{\cdot}_j$ is the mean of variable $\cdot$ for caste group j.

Considering *Tagadhari* workers as the dominant/reference group and *Matwali* and *Pain Nachalne* workers as the non-dominant/comparison groups caste wage differentials among these groups can be decomposed into *explained* and *unexplained*

components by employing the Oaxaca (1973) and Blinder (1973) decomposition methodology. The *explained* component (explained by personal characteristics) corresponds to the mean wage that non-dominant group would receive if they were paid according to dominant groups wage structure. The *unexplained* component, in the other hand, is the difference in wage structure between these two groups evaluated by the mean of non-dominant groups personal characteristics and thus reflects the part of the wage gaps attributable to market discrimination.

In the conventional Oaxaca methodology, the gross difference in mean log wages between the two groups can be decomposed into explained differences in the individual productivity characteristics (i.e. differences in E, S and X) and unexplained differences in the market valuation of such individual productivity characteristics (i.e. differences in β, δ and γ),

$$\begin{aligned} \overline{w}_t - \overline{w}_m &= \beta_t(\overline{E}_t - \overline{E}_m) + (\beta_t - \beta_m)\overline{E}_m \\ &+ \delta_t(\overline{S}_t - \overline{S}_m) + (\delta_t - \delta_m)\overline{S}_m \\ &+ \gamma_t(\overline{X}_t - \overline{X}_m) + (\gamma_t - \gamma_m)\overline{X}_m, \end{aligned} \tag{3.11}$$

$$\begin{aligned} \overline{w}_t - \overline{w}_p &= \beta_t(\overline{E}_t - \overline{E}_p) + (\beta_t - \beta_p)\overline{E}_p \\ &+ \delta_t(\overline{S}_t - \overline{S}_p) + (\delta_t - \delta_p)\overline{S}_p \\ &+ \gamma_t(\overline{X}_t - \overline{X}_p) + (\gamma_t - \gamma_p)\overline{X}_p. \end{aligned} \tag{3.12}$$

This study estimates the full decomposition model in equations (3.11) and (3.12) to evaluate the sources of caste wage differentials. For each decomposition, the first term denotes the wage difference attributable to the difference in observable characteristics between the two groups evaluated according to the dominant group's wage structure. The second term represents the wage difference because of differences

in the wage structure between the two groups, evaluated at the mean level of the comparison groups. The former terms represent the explained components of the wage differential whereas the latter terms are the unexplained components. These are also known respectively as *pre-market* discrimination and *current market* market discrimination.

The decomposition in E analyzes differences in education, which in the traditional Oaxaca decomposition is the main component of human capital.

The decomposition in S shows group differences in access to better jobs and this is the main contribution of this study. As argued in Banerjee and Knight(1985), the choice of occupation can influence the wage a worker receives and that this is important for the rigid caste structure in India. Their methodology isolates the effect of productivity characteristics and occupational distribution on wages (see also Hinks and Wastson, 2001, for a related analysis). As it is argued above, access to jobs in medium and large firms can play a considerable role in producing wage differentials across groups of workers and this is particularly important for developing countries in which the average firm size is smaller than in developed countries. In order to evaluate the effect of occupation and firm size on caste wage differentials three models are considered . First, it applies the occupation decomposition, $S = \{occupation\}$; second, only apply the firm size decomposition, $S = \{firm\ size\}$; and finally, it considers decomposing the full interaction between occupation and firm size, $S = \{occupation \times firm\ size\}$. These models are referred as *Occupational*, *Firm size* and *Interaction* decomposition models, respectively.

Finally, the decomposition in X studies other characteristics such as industry, rural/urban or regional distribution of workers cannot be ruled out while estimating the sources of wage differentials across castes.

3.6 Results

3.6.1 Analysis of access to large firms and occupations

3.6.1.1 Access to large firms

Since this studies argues that caste-differences in access to large firms could be an important factor to produce caste wage differentials a multivariate probit analysis is carried out to see whether these differences exist in Nepalese labour market. Results are presented in Table 3.6. Marginal effects are reported instead of coefficients. Columns 1 and 2 in this Table presents marginal effects with and without interaction terms (caste and education variables), respectively for 2003. Same results are listed in columns 3 and 4 for 2010. This probit model has a dependent variable carrying value one if firm size is categorised as large and zero otherwise. Independent variables are same as used in other models.

Results show that educational attainment positively affects access to larger firm in both years. Land holding and rural dummy variables have impact on access to large firms only in 2003. Land holding has a positive coefficient while the rural dummy coefficient is negative. It shows that individuals from rich households are likely to get jobs in large former whereas those born in rural areas are less likely relative to urban-born individuals. However, both coefficients are not significant in the latter period which might indicate the expansion of larger firms in rural areas over time. Similarly, regional variables, central and western regions, positively affect an individual's probability of being employed in large firms in the first period of analysis. These results were expected since both regions are more urbanized compared to far western region (reference category) and most of the industries are located in these regions. In the latter period, central region continues to show positive and statistically significant coefficients while coefficients for western region is not

statistically significant. Eastern region shows positive and statistically significant coefficients in this period. This indicates that there is a significant shift in industrial establishments across regions over the period of analysis.

Occupational and industry-types distribution of workers are not generally seen as important factors for being associated with large firms. Finally, coefficients for caste dummy variables show that disadvantaged castes are less likely to be employed in larger firms relative to the dominant caste in 2003. In the latter period both caste variables are negative. However, the *Matwali* coefficient is not statistically significant. It indicates that the lowest caste (*Pani Nachalne*) continues to be constrained to work in small firms whereas there might be some improvements in case of intermediate (*Matwali*) caste. Moreover, when interaction terms (interaction between castes and education) are included as additional explanatory variables caste coefficients remain negative and statistically significant in 2003 but not in 2010. It implies that low caste worker's limited access to larger firm has been perpetuated because of the low level of human capital endowment in them.

3.6.1.2 Access to occupations

As in access to large firm, low castes are also seen to be engaged in low paid manual jobs especially in 2003. For instance, marginal effects reported in Table 3.7 show that both the *Matwali* and *Pani Nachalne* castes are less likely to have professional occupation compared to the dominant caste. Both lower castes are likely to have occupations categorised as service, agri-worker and skilled. However, caste coefficients associated with service and agri-worker occupations are statistically insignificant. In the second period, the *Matwali* coefficients continue to show negative signs for professional and clerical occupations but positive signs for agri-worker and skilled occupations. The *Matwali* coefficients for clerical, sales and skilled are statistically

significant. The former two coefficients are negative while the latter has a positive sign. It shows that the *Matwali* workers are less likely to have clerical and sales occupations but more likely to work as skilled worker relative to workers from the *Tagadhari* group which is consistent to the results obtained in 2003. None of the coefficients associated with the *Pani Nachalne* caste appear statistically significant in this period. It was not expected since the lowest caste is more likely to face occupational constraints relative to intermediate caste. However, investigating the underlying factors for these contradicting results is beyond the scope of this study. Overall, these results indicate that impact of occupational distribution on caste wage differentials is decreasing over time.

3.6.2 Analysis of sources of wage differentials

3.6.2.1 Baseline regression analysis

Regression analysis was carried out to estimate the underlying wage equations for each sample period. The estimates are listed in Tables 3.8 for 2003 and 3.9 for 2010. Columns 1, 2 and 3 report results of separate regressions for each of the three castes, followed by the pooled sample results in column 4 with caste dummy variables, where the *Tagadhari* caste represents the reference caste.

Returns to education for the pooled sample are positive, increasing with time, 0.017 (significant at 5%) and 0.066 (significant at 1% level) in 2003 and 2010, respectively. However, they vary considerably across caste groups. In 2003, the *Tagadhari* caste has positive and significant returns, *Matwali* depicts a negative but statistically not significant education coefficient. The *Pani Nachalne* has a positive but statistically insignificant coefficient in this period. In 2010, these coefficients increased markedly for each group and are statistically significant. The *Tagadhari* and *Pani Nachalne* have identical return to education. However, these vary in the

level of significance. The *Tagadhari* coefficient is significant at 1% level while the *Pani Nachalne* coefficient is signficant only at 10% level. The *Matwali* has the lowest, but still fair, return to education which is significant at 1% level.

Firm size plays a crucial role in determining wages in the *Tagadhari* and *Matwali* sub-samples. For example, in 2003, those belonging to the *Tagadhari* group and working in medium-sized and large-sized firm were likely to earn a premium of respectively 35.4% and 57.3% compared to those working in small firms. The same measures account for 56.9% and 54.3% for the *Matwali* sub-sample. Firm size coefficients other than the medium firm in the *Matwali* sub-sample are similar in the latter period. These coefficients are statistically not significant in the *Pani Nachalne* sub-sample.

The results for occupational effects (with reference group = unskilled workers) show mixed significance across sub-samples. For instance, professional, clerical and skilled occupations are the main contributors of the *Tagadhari* worker's wage in 2003. Occupational categories other than professional and sales do not show any significant impact on *Matwali* worker's wages in this period. None of the occupation coefficients are found statistically significant in the *Pani Nachalne* sub-sample.

In the second period, professional occupation continues to have a positive impact on the *Tagadhari* worker's wage whereas professional, clerical and skilled occupations seem to have positive impact on the *Matwali* worker's wage. As in 2003, none of the occupations seem to have significant impact on wage earning by the *Pani Nachalne* workers.

Industry-type effects (with reference group = Agriculture) are not consistent across sub-samples and reflect variability in the base category.

In the pooled regression using caste dummies in column 4, the coefficients on the dummies are negative for both castes in 2003. However, the *Matwali* coefficient is

not statistically significant in this period. In contrast, both caste dummy coefficients became positive although still not significant in 2010. This shows that in order to explore the sources of caste wage differentials, the Oaxaca decomposition model is necessary.

3.6.2.2 Decomposition results

Three different decomposition models are employed to study the sources of wage differentials. These models are hereafter referred as the *Occupational*, *Firm size* and *Interaction* decomposition models. Each model consists of three components; namely (1) explained and unexplained wage differences attributable to differences in education endowments, (2) explained and unexplained wage differences attributable to differences in job characteristics (firm size and/or occupation), (3) explained and unexplained wage differences attributable to differences in other variables including the constant term.

The results are presented in Tables 3.10 and 3.11 for the years 2003 and 2010, respectively. These tables report only the summary results. Detailed decomposition results are not presented in order to save space but can be provided upon request.

The decomposition results show that wage gaps attributable to differences in human capital endowments (i.e. education, explained, $\beta_t(\overline{E}_t - \overline{E}_m)$ and $\beta_t(\overline{E}_t - \overline{E}_p)$), generally considered as being the main source of wage gaps among workers, explains less than half of the wage differentials in 2003 but more than three-fourths in 2010. For 2003 and for the *Tagadhari - Matwali* wage differential, the *Occupational* model shows that differences in education endowments are 0.060 out of a total wage gap of 0.299, and this corresponds to 0.096 and 0.057 for for the *Firm size* and *Interaction* models. For the *Tagadhari - Pani Nachalne* wage differential, the *Occupational* model shows that differences in education endowments are 0.080 out

of 0.493, and this corresponds to 0.128 and 0.076 for the *Firm size* and *Interaction* models. In 2010, the *Tagadhari - Matwali* wage differential decreases to 0.199, and this is explained by differences in education endowments by 0.179, 0.213, 0.150 for the *Occupational*, *Firm size* and *Interaction* decomposition models, respectively. Moreover the *Tagadhari - Pani Nachalne* wage differential is 0.489 in 2010, and this is explained by differences in education endowments by 0.380, 0.454, 0.319 for the *Occupational*, *Firm size* and *Interaction* decomposition models, respectively.

The wage gaps arising from differences in job characteristics (i.e. job, explained, $\delta_t(\overline{S}_t - \overline{S}_m)$ and $\delta_t(\overline{S}_t - \overline{S}_p)$) are statistically significant, and they show a consistent positive effect. The results show that the largest effect is obtained when using the *Interaction* decomposition model. Overall, this shows that access to jobs in better occupations and higher paying firms plays a non-trivial part in explaining the wage gaps across castes. In 2003, for the *Tagadhari - Matwali* wage differential, differences in occupation explain a gap of 0.127, differences in firm size explain 0.077 and the interaction of the two 0.180 (out of 0.299); while for the *Tagadhari - Pani Nachalne* wage differential, each model explains 0.128, 0.063, 0.191 (out of 0.493), respectively. In 2010, for the *Tagadhari - Matwali* wage differential, differences in occupation explain a gap of 0.041, differences in firm size explain 0.032 and the interaction of the two 0.084 (out of 0.199); while for the *Tagadhari - Pani Nachalne* wage differential, each model explains 0.088, 0.078, 0.227 (out of 0.489), respectively.

The differences in endowments in variables other than education, occupation and firm size (i.e. others, explained, $\gamma_t(\overline{X}_t - \overline{X}_m)$ and $\gamma_t(\overline{X}_t - \overline{X}_p)$) generally appear as statistically insignificant. Moreover, the unexplained differences in wage gaps attributable to education (i.e. differences in returns to education), job characteristics (occupation and/or firm size), and other components are in general not statistically significant, although some of them are large in magnitude. Note that the latter

contains industry as one component which preliminary estimations show it is not relevant for the decomposition.

One important point to arise from this analysis is that the *Tagadhari - Matwali* wage differential decreased in 2010 whereas the *Tagadhari - Pani Nachalne* wage differential remained constant. The underlying reason could be that there is a slightly reduction in the gaps in human capital endowment in the former comparison group which has been widened in the case of the latter group. The *Matwali* group have improve their access to better jobs with a relative improvement in educational attainment in the latter period. For instance, *Interaction* decomposition results shows that the job-explained component of the *Tagadhari- Matwali* wage differential has decreased to 0.084 in 2010 relative to 0.180 in 2003 while it has increased in the case of *Tagadhari- Pani Nachalne* wage differential. This indicates that although government introduced a policy of affirmative action providing *quotas* in public sector jobs, the *Pani Nachalne* group might not have been able to take this advantage because of a lack of minimum level of education required for public sector jobs.

3.6.2.3 Robustness: Imputation of missing firm size

Preceding analysis was restricted to a subset of workers who had explicitly reported the firm size of their employer. This exclusion had resulted in a higher proportion of workers in the professional and clerical occupations than in the overall sample. It could therefore be suspected that the estimated decomposition results may be attributable to group differences in access to white collar jobs rather than group differences in access to larger firms. Thus this study proposes another firm size measurement that might still suffer from measurement error but that serves to evaluate the robustness of the previous results. Note that both, previous and new, firm size variables are (imperfect) proxies for the quality of the firm and the job.

In order to test for this possible bias, an extended sample is constructed by imputing a large firm size when missing for certain occupations where the size can be detectable from the work description reported in the survey questionnaire but imputing the rest to small firm. Work descriptions given by production/operation department managers, architect, engineers, nursing/midwifery professionals, primary and secondary education teachers, other teaching professionals, business professionals, computer technicians, optical/electronic equipment operators, modern health associates, administrative personal, secretaries/clerks, library/mail clerks, cashier/tellers clerks, client information clerks, travel attendants, housekeeping and restaurant workers are considered as working for the large firm. The rest of workers with missing firm size are imputed as small firm, except for agricultural, fishery, brick/glass workers and porters for which firm size cannot be clearly assigned and they are therefore excluded from the imputation exercise.

Table 3.12 reports the original and imputed firm size distribution. It should be noted that the imputation exercise increases mostly those assigned to small firms. This imputation leads to a significant increment in the sample size (from 785 to 1357 in 2003 and from 834 to 1110 in 2010) and a reduction in the proportion of white collar jobs. The proportion of professional and clerical workers is reduced to 23.37% and 7.59% from 38.30% and 12.08% in 2003, respectively, and to 19.91% and 14.59% from 23.86% and 18.71% in 2010. Tables 3.13 and 3.14 presents the distribution of male wage workers by occupation and industry, before and after the imputation exercise.

Decomposition results for the extended sample are listed in Tables 3.15 and 3.16. If the difference in access to white collar occupations was driving the baseline results is valid, then it is expected that the explained components of access to jobs will be smaller in the extended sample than in the baseline sample, particularly for 2003

where the proportion of white collars jobs has been significantly reduced in the extended sample. In 2003, the results for the Job-Explained component increases while the Education-Explained component is slightly reduced. For instance, in the *Interaction* model, the Job-Explained increases to 0.211 from 0.180 in the *Matwali* and to 0.225 from 0.191 for the *Pani Nachalne* groups. In 2010, on the contrary, the Job-Explained component decreases although the *Interaction* model still continues to have the largest effect. Overall the results are qualitatively similar to those of the original sample, and thus, they confirm that access to jobs in larger firms play an important role in explaining caste wage differentials.

3.7 Conclusion

This study investigated the sources of caste wage differentials in Nepal by expanding the conventional Oaxaca methodology to include both occupational and firm size effects. The study covered two different surveys over a time span of seven years and included a period of radical political change in Nepal. These changes led to the adoption of policies designed to equalize labour market outcomes in the country.

It finds that caste wage inequality is present in the Nepalese labour market in both years, and it has remained constant between the two periods particularly for the lowest caste. At the same time, results indicate that differences in human capital endowments are important for explaining wage inequality, but so are occupational and firm size effects, especially when the the latter two are taken together. Within the components of discrimination that are related to access to better jobs results indicate that such access continues to exist for reasons other than differences in human capital for both *Matwali* and *Pani Nachalne* disadvantaged groups. This suggests that discriminatory behaviour by employers continues to exist in Nepal.

Overall, this suggests that the government's policy of trying to reverse historical

caste labour market discrimination by imposing *quotas* in public sector employment has not been enough to overcome other barriers that prevent under-privileged workers from accessing such jobs.

Table 3.1: Nepal social hierarchy: 1854

Hierarchy	Habitat	Belief/Religion
A.**Water acceptable(pure)**		
1.*Tagadhari:* Wearer of the sacred thread		
"Upper Caste" (Brahmin)	Hills	Hinduism
" Upper caste" (Madhesi)	Tarai	Hinduism
" Upper Caste" (Newar)	Kathmandu Valley	Hindusim
2. *Matwali:*Alcohol drinkers(non-enslavable)		
Gurung, Magar, Sunuwar	Hills	Tribal / Shamanism
Thakali, Rai, Limbu	Hills	Tribal / Shamanism
Newar	Kathmandu Valley	Buddhism
3. *Matwali:*Alcohol drinkers(enslavable)		
Bhote(Tamang)	Mountain/Hills	Buddhisim
Gharti,Chepang, Hayu	Hills	
Kumal , Tharu	Inner Tarai	Animism
B. **Water unacceptable (impure)**		
1. *Pani Nachalne*:Touchable		
Dhobi, Kasai, Kusule, Kalu	Kathmandu Valley	Hinduism
Musalman	Tarai	Islam
Mlechha(Foreigner)	Europe	Christianity etc.
2. *Pani Nachalne*: Untouchable(achhut)		
Badi, Damai ,Gaine	Hill	Hinduism
Kadara, Kami, Sarki(Parbatiya)	Hills	Hinduism
Chhyame, Pode (Newar)	Kathmandu Valley	Hinduism

Source: Adapted from Bennett et al.(2008).

Table 3.2: Distribution of workers by age categories

Age group	*Year: 2003*	*Year: 2010*
19-29	34.35%	35.43%
30-39	30.79%	26.96%
40-49	20.74%	24.43%
50-59	14.12%	13.18%

Table 3.3: Descriptive statistics: 2003

Variables	Total	*Tagadhari*	*Matwali*	*Pani Nachalne*
Caste	1.00	.707(.016)	.192(.014)	.099(.010)
Lhwage	3.34(.033)	3.45(.039)	3.15(.073)	2.96(.099)
Education	7.78(.172)	8.53(.201)	6.24(.361)	5.5(.557)
Experience	20.69(.411)	20.22(.475)	21.38(.946)	22.65(1.50)
Experience^2	560.66(20.32)	534.35(22.81)	591.60(48.35)	687.62(80.33)
Married	.825(.013)	.810(.016)	.880(.026)	.820(.043)
Rural	.798(.014)	.761(.018)	.934(.020)	.794(.046)
Lnholding('00000)	7.34(.733)	8.44(.994)	6.08(1.04)	2.01(.308)
Small firm	.059(.008)	.043(.008)	.106(.025)	.077(.030)
Medium firm	.419(.017)	.383(.020)	.497(.040)	.526(.056)
Large firm	.522 (.017)	.574(.021)	.397(.039)	.397(.055)
Eastern	.121(.011)	.096(.012)	.139(.028)	.253(.049)
Central	.421(.017)	.447(.021)	.374(.039)	.333(.053)
Western	.136(.012)	.125(.014)	.189(.031)	.116(.036)
Mid-western	.070(.009)	.066(.010)	.083(.022)	.077(.030)
Far-western	.046(.007)	.047(.008)	.063(.019)	-
Abroad	.206(.014)	.219(.017)	.152(.028)	.221(.046)
Unskilled	.184(.013)	.161(.015)	.278(.036)	.167(.042)
Professional	.386(.017)	.451(.021)	.245(.035)	.192(.044)
Clerical	.122(.011)	.133(.014)	.073(.021)	.128(.038)
Service	.057(.008)	.045(.008)	.086(.022)	.090(.032)
Sales	.047(.007)	.054(.009)	.026(.013)	.038(.021)
Agri-worker	.025(.005)	.014(.004)	.046(.017)	.064(.027)
Skilled	.179(.013)	.142(.014)	.246(.035)	.321(.053)
Agriculture	.022(.005)	.019(.005)	.026(.013)	.026(.018)
Mining	.014(.004)	.013(.004)	.020(.011)	.013(.012)
Manufacturing	.193(.014)	.152(.015)	.238(.034)	.397(.055)
Construction	.034(.006)	.029(.007)	.066(.020)	.012(.012)
Trade	.093(.010)	.107(.013)	.060(.019)	.064(.027)
FRE	.034(.006)	.045(.008)	.013(.009)	-
Servicesec	.451(.017)	.471(.021)	.444(.040)	.321(.053)
Others	.159(.013)	.164(.015)	.133(.027)	.167(.042)
Obs.	785	554	153	78

Notes: Standard errors in parentheses.
"-" indicates no observations.

Table 3.4: Descriptive statistics: 2010

Variables	Total	*Tagadhari*	*Matwali*	*Pani Nachalne*
Caste	1.00	.713(.015)	.214(.014)	.073(.009)
Lhwage	3.89(.029)	3.96(.034)	3.76(.059)	3.47(.102)
Education	9.88(.129)	10.66(.131)	8.56(.304)	6.21(.573)
Experience	19.56(.392)	19.27(.457)	20.77(.892)	18.91(1.48)
Experience^2	510.11(17.95)	495.00(20.51)	569.31(42.33)	487.58(70.51)
Married	.792(.014)	.790(.016)	.810(.029)	.766(.055)
Rural	.731(.015)	.710(.018)	.815(.029)	.786(.052)
Lnholding('00000)	29.92(3.98)	36.80(5.15)	14.74(6.84)	6.00(1.95)
Small firm	.030(.006)	.023(.006)	.052(.016)	.067(.032)
Medium firm	.332(.016)	.290(.018)	.339(.035)	.617(.063)
Large firm	.638(.016)	.685(.019)	.609(.037)	.316(.060)
Eastern	.105(.010)	.094(.012)	.126(.025)	.133(.044)
Central	.608(.016)	.652(.019)	.551(.037)	.350(.062)
Western	.157(.012)	.148(.014)	.167(.028)	.217(.053)
Mid-western	.073(.009)	.064(.010)	.075(.019)	.150(.046)
Far-western	.038(.006)	.027(.006)	.052(.016)	.100(.008)
Abroad	.019(.004)	.013(.004)	.029(.012)	.050(.028)
Unskilled	.084(.009)	.072(.010)	.126(.025)	.083(.035)
Professional	.239(.014)	.283(.018)	.149(.027)	.067(.032)
Clerical	.191(.013)	.224(.017)	.086(.021)	.166(.048)
Service	.127(.011)	.115(.013)	.121(.024)	.267(.057)
Sales	.066(.008)	.071(.010)	.046(.015)	.083(.035)
Agri-worker	.008(.003)	.001(.001)	.023(.011)	.033(.023)
Skilled	.282(.015)	.231(.017)	.448(.037)	.300(.059)
Agriculture	.007(.002)	.008(.003)	.005(.005)	-
Mining	.008(.003)	.008(.003)	.011(.008)	-
Manufacturing	.129(.011)	.106(.012)	.149(.027)	.300(.059)
Construction	.035(.006)	.027(.006)	.052(.016)	.067(.032)
Trade	.079(.009)	.081(.011)	.051(.016)	.133(.044)
Servicesec	.193(.013)	.179(.015)	.247(.032)	.183(.050)
FRE	.065(.008)	.074(.010)	.057(.017)	-
Others	.481(.017)	.515(.020)	.425(.037)	.317(.060)
Obs.	834	594	179	61

Notes: Standard errors in parentheses.
"-" indicates no observations.

Table 3.5: Wages by occupation and firm size

	Year: 2003			Year: 2010		
Occupation	Small firm	Medium firm	Large firm	Small firm	Medium firm	Large firm
Unskilled	2.26(0.772)	2.90(0.941)	3.09(0.608)	3.26(.769)	3.37(0.617)	3.64(0.715)
Professional	3.08(1.50)	3.48(0.907)	3.91(0.874)	4.50(1.27)	4.50(0.936)	4.52(0.762)
Clerical	2.16(1.26)	3.14(0.628)	3.78(0.761)	-	3.86(0.846)	4.02(0.638)
Service	2.65(0.951)	3.02(1.06)	3.19(0.393)	-	3.19(0.704)	3.83(0.723)
Sales	2.59(0.260)	2.65(0.868)	3.13(.291)	3.09(0.580)	3.23(0.460)	3.66(0.640)
Agri-worker	3.06(.659)	3.14(1.03)	3.67(0.792)	3.17(2.52)	-	3.41(0.431)
Skilled	2.77(0.490)	3.14(0.897)	3.15(0.793)	3.07(0.430)	3.28(0.720)	3.96(0.650)

Note: Standard deviation in parentheses.
"-" indicates no observations.

Table 3.6: Probit regression results

Dependent variable: dummy=1 for large firm, 0 otherwise				
Variables	*Year: 2003*		*Year: 2010*	
	1	2	3	4
Education	.009**(.004)	.007(.005)	.035***(.006)	.040***(.008)
Experience	.011*(.007)	.013*(.007)	.006(.006)	.006(.006)
Experience2	-.000(.000)	-.000(.000)	.000(.000)	.000(.000)
Married	-.053(.061)	-.052(.062)	.039(.058)	.041(.058)
Rural	-.158**(.069)	-.164**(.069)	-.031(.042)	.032(.042)
Lnholding	.365***(.120)	.373***(.120)	-.011(.015)	-.011(.015)
Eastern	.075(.102)	.066(.104)	.222***(.057)	.218***(.057)
Central	.195**(.089)	.194**(.089)	.123*(.059)	.116*(.059)
Western	.201**(.091)	.194**(.092)	-.112(.090)	-.113(.090)
Mid-western	.088(.109)	.092(.108)	-.026(.103)	-.023(.104)
Abroad	.076(.096)	.073(.096)	.092(.120)	.079(.122)
Professional	.098*(.060)	.101*(.060)	-.156*(.086)	-.165*(.087)
Clerical	.039(.074)	-.040(.074)	-.031(.082)	-.033(.082)
Service	.112(.086)	.124(.087)	.009(.085)	.005(.085)
Sales	-.382***(.076)	-.382***(.076)	-.483***(.113)	-.485***(.113)
Agri-worker	-.183(.125)	-.170***(.127)	.272*(.085)	.251*(.100)
Skilled	-.009(.063)	-.002(.063)	-.048(.073)	-.050(.073)
Mining	.310(.157)	.324*(.148)	.136(.223)	.135(.222)
Manufacturing	.052(.150)	.065(.150)	.162(.145)	.153(.145)
Construction	-.214(.157)	-.204(.159)	.120(.163)	.119(.160)
Trade	-.134(.157)	-.129(.157)	.146(.162)	.140(.161)
FRE	-.013(.182)	-.002(.181)	.018(.176)	.014(.173)
Servicesec	-.192(.143)	-.190(.142)	.142(.149)	.134(.150)
Others	.017(.149)	.025(.148)	.221(.168)	.216(.165)
*Matwali*Education*	-	.005(.011)	-	-.009(.011)
*Pani Nachalne*Education*	-	.021*(.013)	-	-.016(.015)
Matwali	-.154***(.049)	-.196**(.083)	-.047(.047)	.037(.109)
Pani Nachalne	-.154**(.063)	-.270***(.092)	-.207***(.080)	-.073(.138)
Pseudo R^2	.1156	.1179	1629	.1640
Log likelihood ratio	-480.75	-479.50	-457.02	-456.40
Obs.	785	785	834	834

Notes: Robust standard errors in parentheses. * significant at 10%, ** significant at 5% and *** significant at 1%.

-Far-western, Unskilled, Agricultural and *Tagadhari* are omitted categories for region, occupation, industry-type and caste dummy variables, respectively.

-These coefficients correspond to marginal effects.

Table 3.7: Multinomial regression results

Dependent variable: Occupational categorical variable						
Year: 2003						
	Professional	*Clerical*	*Service*	*Sales*	*Agri-worker*	*Skilled*
Education	.034***(.004)	.009***(.003)	-.003***(.001)	-.001(.001)	-.003***(.001)	-.012***(.002)
Experience	-.003**(.002)	-.001(.001)	.000(.000)	.000(.000)	.001(.001)	-.001(.001)
Rural	.015(.047)	-.053(.047)	.079**(.036)	-.002(.020)	.021**(.009)	-.050(.039)
Lnholding	.439***(.145)	.269***(.071)	-.382***(.123)	-.073(.122)	.039(.028)	-.73*(.159)
Matwali	-.164***(.044)	-.030(.033)	.030(.022)	-.029*(.017)	.021(.016)	.119***(.043)
Pani Nachalne	-.185***(.056)	.036(.048)	.020(.027)	-.016(.023)	.040(.029)	.167***(.058)
Log likelihood ratio			-1154.62			
Obs.			785			
Year: 2010						
Education	.046***(.009)	.033***(.007)	-.025***(.004)	-.003(.003)	-.001(.001)	-.034***(.007)
Experience	.002***(.000)	.002**(.001)	-.005***(.001)	-.003***(.001)	.000(.000)	.000(.000)
Rural	-.010(.010)	-.045(.029)	.032(.033)	.015(.033)	-.001(.004)	.048(.042)
Lnholding	-.002(.003)	.010(.023)	.014(.024)	.001(.008)	.000(.000)	.013(.016)
Matwali	-.007(.010)	-.154***(.038)	-.028(.033)	-.038*(.023)	.007(.007)	.188***(.047)
Pani Nachalne	.010(.028)	.031(.075)	.048(.055)	.001(.041)	.007(.009)	-.055(.073)
Log likelihood ratio			-1150.02			
Obs.			834			

Notes: Robust standard errors in parentheses. * significant at 10%, ** significant at 5% and *** significant at 1%.
- Unskilled and *Tagadhari* are base categories for occupation and caste variables, respectively.
-These coefficients correspond to marginal effects.

Table 3.8: Regression results: 2003

	Tagadhari	*Matwali*	*Pani Nachalne*	*Dummy*
	1	2	3	4
Education	.025***(.009)	-.017(.024)	.006(.036)	.017**(.007)
Experience	.025*(.015)	.004(.026)	-.038(.036)	.025**(.011)
Experience2	-.000(.000)	.000(.000)	.000(.000)	-.000(.000)
Married	.079(.107)	.384*(.249)	.543**(.248)	.140*(.091)
Rural	.015(.164)	-.095(.264)	-.116(.353)	-.040(.136)
Lnholding	-.121(.133)	-.192(.327)	.244(4.60)	-.144(.123)
Medium firm	.354*(.199)	.569**(.267)	-.413(.430)	.319**(.141)
Large firm	.573***(.198)	.543**(.263)	.291(.476)	.540***(.142)
Eastern	.074(.229)	-.114(.271)	-.357(.271)	-.041(.175)
Central	.275(.195)	.177(.252)	-.359(.369)	.194(.155)
Western	.172(.204)	.916***(.343)	dropped	.291*(.170)
Mid-western	.268(.225)	.109(.295)	-.027(.606)	.210(.178)
Abroad	.218(.205)	.441*(.284)	-.248(.428)	.172(.165)
Professional	.639***(.114)	.667***(.238)	.144(.405)	.668***(.094)
Clerical	.309**(.126)	.104(.331)	.507(.439)	.410***(.111)
Service	.088(.198)	.239(.297)	.035(.465)	.121(.158)
Sales	.082(.215)	.512*(.277)	-.505(.522)	.125(.170)
Agri-worker	.414(.453)	-.255(.355)	-.122(.579)	.382*(.214)
Skilled	.200*(.132)	.196(.193)	.424(.479)	.301***(.100)
Mining	.086(.282)	-.260(.586)	.088(.536)	-.017(.248)
Manufacturing	.231(.277)	-.759**(.327)	.300(.422)	.015(.196)
Construction	.359(.292)	.302(.398)	1.42***(.453)	.240(.215)
Trade	-.074(.283)	-.738**(.356)	.140(.477)	-.208(.204)
FRE	.901***(.303)	-.802(.601)	-	.691***(.247)
Servicesec	.274(.256)	-.311(.339)	-.026(.379)	.088(.184)
Others	.191(.268)	.405(.439)	.628(.573)	.339(.195)
Matwali	-	-	-	-.076(.080)
Pani Nachalne	-	-	-	-.201*(.110)
Constant	1.53***(.406)	2.07***(.551)	2.86***(.813)	1.71***(.303)
R^2	.2717	.3683	.4109	.2613
Obs.	554	153	78	785

Notes: Robust standard errors in parentheses. * significant at 10%, ** significant at 5% and *** significant at 1%.

-Small firm, Far-western, Unskilled, Agricultural and *Tagadhari* are omitted categories for firm size, region, occupation, industry-type and caste dummy variables, respectively.

Table 3.9: Regression results: 2010

	Tagadhari	*Matwali*	*Pani Nachalne*	*Dummy*
	1	2	3	4
Education	.077***(.012)	.057***(.018)	.077*(.041)	.066***(.009)
Experience	.015(.012)	.077***(.021)	.044(.044)	.035***(.010)
Experience2	.000(.000)	-.001***(.000)	-.000(.000)	-.000(.000)
Married	.208*(.144)	-.067(.155)	.166(.308)	.139*(.087)
Rural	-.009(.071)	-.038(.119)	-.153(.326)	-.022(.060)
Lnholding	.024(.023)	-.033*(.022)	-.898(.862)	.010(.020)
Medium firm	.305**(.146)	.304*(.205)	-.258(.780)	.264*(.137)
Large firm	.492***(.149)	.580***(.189)	.150(.825)	.485***(.138)
Eastern	.142(.146)	-.387(.393)	-.233(.676)	.008(.138)
Central	.205*(.121)	-.125(.344)	.282(.674)	.188*(.123)
Western	.016(.148)	-.359(.356)	.019(.667)	-.008(.138)
Mid-western	.260(.197)	-.546(.388)	.409(.678)	.147(.169)
Abroad	.306(.309)	-.913**(.376)	.081(.794)	.008(.212)
Professional	.444***(.138)	.561***(.204)	.372(.790)	.545***(.108)
Clerical	.146(.123)	.306*(.206)	-.217(.685)	.229**(.098)
Service	.055(.136)	.223(.223)	-.326(.501)	.139(.105)
Sales	-.666***(.198)	-.851***(.318)	.290(.710)	-.426**(.171)
Agri-worker	.160(.140)	-.256(.414)	-.460(.538)	-.375(.260)
Skilled	.060(.120)	.323**(.156)	-.134(.561)	.190*(.088)
Mining	-.758**(.347)	.079(.363)	-	-.478(.338)
Manufacturing	-.290(.318)	.534*(.360)	.671(.666)	-.026(.305)
Construction	-.258(.352)	.810**(.353)	.387(.880)	.040(.322)
Trade	.164(.327)	1.05***(.361)	-	.290(.315)
Servicesec	-.462(.314)	.355(.331)	.247(.630)	-.213(.303)
FRE	-.042(.327)	.774**(.360)	-	.190(.315)
Others	-.248(.311)	.624*(.329)	-.118(.496)	-.014(.301)
Matwali	-	-	-	.043(.056)
Pani Nachalne	-	-	-	.064(.113)
Constant	2.18***(.383)	1.49***(.616)	2.24*(1.35)	1.92***(.380)
R^2	.3691	.4951	.3236	.3716
Obs.	594	179	61	834

Notes: Robust standard errors in parentheses. * significant at 10%, ** significant at 5% and *** significant at 1%.

-Small firm, Far-western, Unskilled, Agricultural and *Tagadhari* are omitted categories for firm size, region, occupation, industry-type and caste dummy variables, respectively. However, Trade industry is omitted in the *Pani Nachalne* sub-sample since it has no observation on base category, Agricultural industry.

Table 3.10: Oaxaca decomposition results: 2003

	Tagadhari vs. Matwali						
		Education		*Job*		*Other*	
	Total	**Explained**	**Unexplained**	**Explained**	**Unexplained**	**Explained**	**Unexplained**
Occupational	.299***	.060**	.198*	.127***	0.029	.016	-.131
	(.089)	(.023)	(.153)	(.036)	(.044)	(.027)	(.200)
Firm size	.299***	.096***	.161	.077***	.059	.041	-.135
	(.086)	(.026)	(.127)	(.025)	(.418)	(.029)	(.297)
Interaction	.299***	.057**	.265*	.180***	.201	.014	-.418
	(.089)	(.023)	(.153)	(.044)	(.259)	(.027)	(.604)
	Tagadhari vs. Pani Nachalne						
Occupational	.493***	.080**	.104	.128***	.178	.041	-.038
	(.118)	(.032)	(.196)	(.042)	(.493)	(.044)	(.378)
Firm size	.493***	.128***	.230*	.063**	.361	.114**	-.403
	(.114)	(.036)	(.152)	(.028)	(95.99)	(.045)	(.419)
Interaction	.493***	.076**	.289*	.191***	-.270	.044	.163
	(.119)	(.031)	(.191)	(.055)	(.478)	(.044)	(.864)

Notes: Standard errors in parentheses. * significant at 10%, ** significant at 5% and *** significant at 1%.

Table 3.11: Oaxaca decomposition results: 2010

	Tagadhari vs. Matwali						
		Education		*Job*		*Other*	
	Total	**Explained**	**Unexplained**	**Explained**	**Unexplained**	**Explained**	**Unexplained**
Occupational	.199***	.179***	.126	.041	-.192	.009	.036
	(.071)	(.037)	(.180)	(.031)	(.108)	(.029)	(.232)
Firm size	.199***	.213***	.202	.032**	-.095	-.003	-.150
	(.070)	.040)	(.166)	(.015)	(.064)	(.030)	.334
Interaction	.199***	.150***	.131	.084*	-.780*	.002	.612
	(.071)	(.034)	(.187)	(.043)	(.501)	(.027)	(.791)
	Tagadhari vs. Pani Nachalne						
Occupational	.489***	.380***	.084	.088*	-.068	.044	-.039
	(.122)	(.071)	(.310)	(.046)	(.071)	(.055)	(.528)
Firm size	.489***	.454***	.027	.078***	.801	.092*	-.963
	(.118)	(.075)	(.246)	(.029)	(1.22)	(.055)	(.592)
Interaction	.489***	.319***	.019	.227***	-.394	.028	.290
	(.128)	(.067)	(.318)	(.064)	(.493)	(.053)	(1.25)

Notes: Standard errors in parentheses. * significant at 10%, ** significant at 5% and *** significant at 1%.

Table 3.12: Firm size distribution (before and after imputation)

	Year: 2003						*Year: 2010*					
	Reported		*Imputed*		*Total*		*Reported*		*Imputed*		*Total*	
	Obs.	*%*	*Obs.*	*%*	*Obs.*	*%*	*Obs.*	*%*	*Obs.*	*%*	*Obs.*	*%*
Small firm	46	5.85	546	95.62	592	43.63	25	2.99	247	89.49	272	24.50
Medium firm	332	42.37	-	-	333	24.54	278	33.29	-	-	278	25.05
Large firm	407	51.78	25	4.38	432	31.83	532	63.72	28	10.14	560	50.45
Total	785	100	571	100	1357	100	834	100	276	100	1110	100

Table 3.13: Distribution of male wage workers by occupation and industry (before and after firm size imputation): 2003

	Reported		*Imputed*		*Total*	
Occupation	*Obs.*	*%*	*Obs.*	*%*	*Obs.*	*%*
Unskilled	145	18.45	99	17.34	244	17.98
Professional	301	38.30	16	2.80	317	23.37
Clerical	95	12.08	8	1.40	103	7.59
Service	46	5.85	28	4.90	74	5.45
Sales	37	4.71	10	1.75	47	3.46
Agri-workers	20	2.54	15	2.63	35	2.58
Skilled	141	18.07	395	69.18	537	39.57
Industry	*Obs.*	*%*	*Obs.*	*%*	*Obs.*	*%*
Agriculture	17	2.16	20	3.50	37	2.73
Mining	11	1.40	6	1.05	17	1.25
Manufacturing	152	19.34	172	30.13	324	23.88
Construction	29	3.69	299	52.37	328	24.17
Trade	71	9.03	16	2.80	87	6.41
FRE	27	3.44	4	0.70	31	2.28
Service sector	351	44.78	32	5.60	384	28.30
Other	127	16.16	22	3.85	149	10.98
Total	785	100	571	100	1357	100

Table 3.14: Distribution of male wage workers by occupation and industry (before and after firm size imputation): 2010

	Reported		*Imputed*		*Total*	
Occupation	*Obs.*	*%*	*Obs.*	*%*	*Obs.*	*%*
Unskilled	75	8.99	2	0.73	77	6.94
Professional	199	23.86	22	7.97	221	19.91
Clerical	156	18.71	6	2.17	162	14.59
Service	107	12.83	58	21.01	165	14.86
Sales	55	6.59	3	1.09	58	5.23
Agri-worker	6	0.72	6	2.17	12	1.08
Skilled	236	28.30	179	64.86	415	37.39
Total	834	100	276	100	110	100
Occupation	*Obs.*	*%*	*Obs.*	*%*	*Obs.*	*%*
Agricultural	7	0.84	18	6.52	25	2.24
Mining	6	0.72	3	1.09	9	0.81
Manufacturing	109	13.07	65	23.55	174	15.68
Construction	28	3.36	117	42.39	145	13.06
Trade	66	7.91	8	2.90	74	6.67
FRE	158	18.94	30	10.87	188	16.94
Service sector	55	6.59	5	1.81	60	5.41
Other	405	48.57	30	10.87	435	39.19
Total	834	100	276	100	1110	100

Table 3.15: Oaxaca decomposition results with imputed firm size: 2003

Tagadhari vs. Matwali							
		Education		*Job*		*Other*	
	Total	**Explained**	**Unexplained**	**Explained**	**Unexplained**	**Explained**	**Unexplained**
Occupational	.268***	.063**	.059	.138***	.011	-.053*	.050
	(.053)	(.028)	(.066)	(.034)	(1.83)	(.035)	(.135)
Firm size	.268***	.107***	.071	.083**	.025	-.032	.014
	(.053)	(.028)	(.063)	(.035)	(.070)	(.038)	(.107)
Interaction	.268***	.044*	.073	.211***	.026	-.086**	.000
	(.053)	(.028)	(.063)	(.047)	(.145)	(.039)	(.199)
Tagadhari vs. Pani Nachalne							
Occupational	.387***	.082**	-.000	.148***	.021	-.039	.175
	(.069)	(.036)	(.070)	(.036)	(.284)	(.042)	(.172)
Firm size	.387***	.140***	.085	.097**	-.031	.010	.086
	(.068)	(.037)	(.065)	(.042)	(.300)	(.043)	(.119)
Interaction	.387***	.057*	.042	.225***	.030	-.069*	.102
	(.070)	(.036)	(.071)	(.052)	(.037)	(.044)	(.224)

Notes: Standard errors in parentheses. * significant at 10%, ** significant at 5% and *** significant at 1%.

Table 3.16: Oaxaca decomposition results with imputed firm size: 2010

Tagadhari vs. Matwali							
		Education		*Job*		*Other*	
	Total	**Explained**	**Unexplained**	**Explained**	**Unexplained**	**Explained**	**Unexplained**
Occupational	.216***	.202***	.029	.041*	-.299	.001	.242
	(.061)	(.035)	(.137)	(.026)	(.327)	(.027)	(.225)
Firm size	.216***	.245***	.137	.013	-.101	-.011	-.067
	(.061)	(.038)	(.130)	(.016)	(.192)	(.029)	(.167)
Interaction	.216***	.179***	.073	.076**	.131	.003	-.246
	(.068)	(.035)	(.143)	(.033)	(.788)	(.026)	(.642)
Tagadhari vs. Pani Nachalne							
Occupational	.489***	.387***	.082	.086***	-.022	.034	-.078
	(.084)	(.061)	(.162)	(.033)	(.021)	(.047)	(.443)
Firm size	.489***	.469***	.081	.043	-.063	.045	-.086
	(.082)	(.062)	(.145)	(.034)	(.110)	(.050)	(.205)
Interaction	.489***	.342***	.063	.150***	1.52	.034	-1.62
	(.086)	(.062)	(.171)	(.054)	(1.98)	(.048)	(1.13)

Notes: Standard errors in parentheses. * significant at 10%, ** significant at 5% and *** significant at 1%.

Chapter 4

Caste Inequity in Health Care Utilisation and The Impact of Caste on Health Outcomes: Evidence from Nepal (joint with professor Gabriel Montes-Rojas and Dr Victoria Serra-Sastre)

4.1 Introduction

Equity is regarded as maintaining fairness or reducing inequality in distribution so as to maintain social justice (Braveman and Gruskin, 2003). Particularly with reference to health sector inequity, inequality in health status, access to and utilisation of health care service across social groups has been extensively examined in the literature. The majority of these studies evaluate such inequities across social groups ranked by income and are largely focused in developed countries (see Schoen and Doty, 2004; van Doorslaer et al., 2000, 2006; Heyden et al., 2003).

Health care systems differ significantly between developing and developed countries, as do the underlying factors of health sector inequity (Mills and Gilson, 1988). One discernible difference in health care systems is accessibility to health care ser-

vices. More unequal access to health care services across social groups can be observed in developing countries than in the developed world. Another difference arises in the composition of different sources of finance in health care services (George and Akiko, 1999). In most developed countries, health care is mainly financed by public sector expenditure, and to a lesser extent by private insurance (there are exceptions, as in the US). In developing countries, the financing of health care expenditure predominantly depends on out-of-pocket (OOP) payments by households. Access to and financing of the health care system in developing countries is therefore likely to favour health sector inequity, possibly driving the poorest segment of the population into higher health risk. Empirical evidence in the developing world is scant (Wagstaff and Doorslaer, 2000).

Health sector inequity in developed and developing countries has mainly been evaluated in terms of the use of health care services, such as physician visits, hospitalization, ambulatory care services and health outcomes across social groups ranked by income and other non-income socioeconomic characteristics such as race and ethnicities. However, scarce attention has been given in empirical studies to inequity across castes despite the fact that a caste system is likely to boost health sector inequities via low social capital in low castes. The caste system imposes a social division of labour based on caste hierarchy that limits low-caste workers to work only in unskilled manual jobs (Banerjee and Knight, 1985). This, in turn, may lead low-caste groups into an inferior socioeconomic status. The available evidence suggests that there is a positive relationship between socioeconomic status, measured by income, and health outcomes both in developed (Deaton, 2003; Wagstaff, 2002; van Doorslaer and Koolman, 2004) and in developing countries (Lordan et al., 2012; Van de Poel and Speybroeck, 2009). Therefore, such social hierarchy can lead low-caste individuals to have worse health outcomes. In addition to income,

other elements of socioeconomic status, such as education and occupation, can also significantly influence individual's health (Winkleby et al., 1992; Adler and Newman, 2002). Deaton (2011) argues that childhood inequalities in social status such as parental education and income are key drivers of adulthood health. The caste system affects all of the above variables. To the best of my knowledge, very few studies have focused on caste inequity in health care utilisation (see Dommaraju et al., 2008; Roy and Howard, 2007; Van de Poel and Speybroeck, 2009) and no study has yet explicitly examined difference in health outcomes across castes.

By acknowledging this gap in the literature of health economics this study first evaluates the caste-inequity in health care utilization, proxied by OOP payments for health care services. In other words, it examines whether there is a caste differentials in OOP payments for health care services which cannot be explained by the differences in need variables (types of illnesses) associated with them. Second, it examines the determining factors of caste disparities in health outcomes as measured by self-assessed-health (SAH).[1] The first effect is evaluated by employing a tobit regression analysis. The second effect is estimated by employing ordered probit and generalized ordered probit models. Two waves of National Living Standard Survey (NLSS) from Nepal for the years 2003 and 2010 are used.

Furthermore, this study covers a period in which various policies were implemented with the aim of providing health care services to the poor. The main initiative includes the introduction of Nepal Health Sector Program-Implementation Plan (NHSP-IP) 2004-2009, which is a strategic plan that outlines Nepal's health policy known as Essential Health Care Service (EHCS), whereby user fees for the poor are abolished and primary health care services are declared universally free. Nepal's recent development plans, which are guiding documents for allocating the annual

[1]Caste categories used in this analysis are identical to one used in previous chapter. Therefore, no detail description is given in this chapter in order to avoid repetition. See section 3.3 for detail.

national budget to different sectoral ministries, have fully incorporated these policy initiatives.[2] Following these development plans, one could expect any caste inequity in health care utilisation to decrease over time had these policy interventions been effective. However, the impact of these policy interventions on caste inequity in health care utilisation in Nepal has not been explored. This study partly fills this gap.

The econometric analysis shows inequity in health care utilisation to exist across castes in Nepal, albeit the evidence shows there has been a decrease in inequity. Furthermore, it shows that low castes are more likely to have lower health outcomes compared to the dominant caste particularly in 2003. However, since caste differences in health outcomes remain in 2010 and the abolishment of user fees does not show a significant impact on caste differences in OOP expenses, it can be argued that accompanying health policies are not effective enough to attain caste equity in health care.

The reminder of this study is organized as follows. Section 4.2 reviews previous literature particularly with reference to inequity in health care utilization and also in difference in health outcomes. Section 4.3 explains the Nepalese health system. Section 4.4 describes the estimation strategy. Data, descriptive statistics and variables specification are presented in Section 4.5 followed by results in Section 4.6. Section 4.7 concludes.

[2]The periodic development plan is an important policy document prepared by the National Planning Commission (NPC) of Nepal for specific periods, usually five years, that sets the priority of the development programs for all sectoral ministries. The annual national budget is allocated in accordance with the priority given by the plan document.

4.2 Literature review

Past few decades have given a tremendous momentum to the literature on health sector inequity. As a result, more than one thousand articles on this subject have been published only in the last two decades (Williams and Cookson, 2000). These studies can broadly be divided into two strands. The first strand of literature evaluates health sector inequity across social groups ranked by income while the second strand focuses on inequity across social classes such as caste, race, ethnicity and gender. These two types of studies are different not only on conceptual grounds but also in methodologies used. The estimation of concentration index is the widely used approach in the former types of studies while regression analysis, decomposition technique or both have been employed frequently in the latter strand of studies.

4.2.1 Socioeconomic inequity in health care utilisation

Despite a substantial upsurge in the empirical studies related to socioeconomic inequity in health care utilisation these are mainly focused in developed counties. A study carried out in OECD countries reveals a pro-rich inequity in health care utilisation to exist virtually in all these countries (van Doorslaer and Wagstaff, 1992). This study defines equity in health care utilisation if the people with identical needs for health care are treated alike.

Chen and Escarce (2004) examine income related inequality in medical expenditure in the United States of America (USA) controlling for need variables proxied by age. Income related inequality is assessed in four different age categories: all adults, working-age adult, seniors and children (with age from 5 to 17 years). Estimated concentration index of inequality for the same age categories was found as .087, .099, .147 and .067, respectively. These results indicate that income related inequity in health care expenditure exists in the USA which favours rich people. The

magnitude of inequity is the highest among seniors relative to other age categories in comparison.

An analysis carried out in Norway by Grytten et al. (1995) contradicts with the conclusion of studies in OECD and in the USA. Looking at the services provided by primary care physicians, this study reveals no significant income effect on its utilisation. This indicates the presence of equality in health care utilisation across income groups. However, these authors acknowledge the possibility of control in the supply of services that might have limited the consumption of health care services for those who wished and were able to afford more amount and high-quality health care services.

Evidence from Asian regions, with some variation in magnitudes, supports the finding of above studies. Lu et al. (2007) investigate horizontal inequity (HI) in health care utilisation in three high-income Asian countries: South Korea, Taiwan and Hong Kong. In line with other studies, this study estimates concentration indexes for utilisation of (visit to) five different categories of health care providers: western doctor, licensed traditional medicine practitioner (LTMP), dental practitioners, emergency room (ER) visits and inpatient admissions. South Korea is shown to maintain equity in health care utilisation while Taiwan and Hong Kong depict a slightly pro-rich bias, particularly in the use of outpatient service and visit to western doctor, respectively.

Other authors have carried out comparative studies in health sector inequity over time. Honda and Ohkusa (2003) examine health inequity employing Comprehensive Survey of Living Condition in Japan (CSLCJ) for the periods of 1992, 1995 and 1998. This study estimates concentration indexes for health care utilisation proxied by physician visits. They find a pro-rich inequity in health care utilisation before 1995 but not in the latter period. A similar study evaluates the change in equality

in health and health care utilisation between urban and rural Chinese during the period of economic transition (transition from command economy to market economy) employing national survey data over the period of 1985 to 1993 (Liu et al.,1999). In contrast to Honda and Ohkusa (2003) this study indicated a widening trend in inequality in health status between urban and rural Chinese which is correlated with increasing gaps in income and health care utilisation between them.

A community environment is also found to influence an individual's access to health care utilisation. A study carried out by Kirby and Keneda (2005) demonstrates that living in a socioeconomically disadvantaged neighborhood undermines an individual's probability of using health care and preventive services but increases the likelihood of having unmet medical need. These authors argue that when disadvantaged groups of people reside densely in a particular area "disadvantage" itself becomes a community specific characteristics to which they describe as "emergent characteristics". Similar evidence were found in a study carried out by Deaton and Lubotsky (2003). This study investigates the effect of racial composition on mortality rate across cities and states in the Unites States. Their results do not show any positive influence of income inequality on mortality rate conditional on racial composition. Therefore, these authors conclude that a positive correlation between income inequality and mortality rate was confounded by the racial composition.

Empirical studies in inequality in health care utilisation are limited in the context of developing or low-income countries (Wagstaff and Doorslaer, 2000). However, available studies do show a positive correlationship between socioeconomic variables and the utilisation of health care services. Makinen et al. (2000) evaluate the probability of obtaining care and receiving medicine in eight developing and transitional countries (Burkina Faso, Guatemala, Kazakhstan, Kyrgyzstan, Paraguay, South Africa, Thailand and Zambia). They find a pro-rich bias in receiving care.

However, this study did not find wealthy people to utilize private health care services.

Another study compares the effect of socioeconomic characteristics and access factor to health care providers on the use of two types of preventive services, immunisation and physical examination, among pre-school children from low-income urban areas in the USA (Wan and Gray, 1978). Access factor is defined as the source of regular use of care and health insurance status associated with children. Utilising data from community health survey this study concludes that the impact of social and access factors on utilisation of both types of services were similar. This study reveals a significant difference in the impact of demographic variables (age, sex and race) on the use of these services. Demographic variables depicted a stronger influence on the immunisation utilisation but only a moderate impact was found with reference to physical examination.

The second strand of empirical studies focuses on inequity in health care utilisation by other non-income socioeconomic characteristics such as gender race and ethnicities. The majority of these studies have focused on racial and ethnic differential in health care utilisation particularly with reference to inequity in the USA.

van Doorslaer and Wagstaff (1992) examine the difference in the amount of physician visit across Hispanic, black non-Hispanic and white non-Hispanic children in USA and evaluate its determinants employing Health Interview Survey (HIS), 1979. This study finds a significant difference in physician visit across these social groups. Additionally, it suggests that financial circumstances play an important role in minority groups when it comes to whether or not visit a physician. In addition, parental perception on health care need for their children influences this decision in case of dominant groups.

Evidence from empirical works bolster that difference in health care utilisation

across social groups is persistent in the USA. Harris et al. (2005) examine the difference in the use of mental health care across nine racial/ethnic groups: White, African American, American Indian / Alaskan native, Asian, Mexican, Central and South American, Puerto Rican and other Hispanic-Latino. Using data from National Survey on Drug Use and Health (NSDUH) for the period of 2001-2003, they find that minority groups (African American, Asian, Mexican, Central and South American, and other Hispanic-Latino) utilise less mental health care services relative to the non-Latino white. Cook et al. (2010) provide further evidence on racial differences in access to mental health services and total mental health expenditure in the USA. Black and Latino-white have less access to mental health care and thus they incur lower mental health care expenditure relative to non-Latino white.

Some authors have investigated whether residential status affects disadvantaged and dominant groups differently with respect to health care utilisation. Berdahl et al. (2007) examine the access to health care services by non-Latino white (Mexican) and Latino white between the non-metro and metro areas of residence in the USA. Two types of care services as measured by physician visits and ambulatory care have been examined in this study. Their results show Mexican being disadvantaged in both measures of care relative to non-Latino white in both residential areas. However, disparities in access to health care between these two groups were greater in non-metro areas.

D'Anna et al. (2010) investigate the effects of perceived discrimination in receiving care and examine its relationship with racial/ethnic disparities in self-reported physical and mental health status employing data from Californian Health Interview Survey (CHIS). Two types of discrimination; discrimination based on racial/ethnic /language (accent) and other discrimination are considered to examine such effects. Findings from this study show negative effects of the former type of discrimination on

self-related health status and it positively influences the functional limitation caused by physical and mental problem in discriminated-against groups. However, these effects vary across race/ethnic groups, gender and socioeconomic position (SEP). They find that the relationship between racial /ethnic minority status and poor health decreases as the level of SES increases.

4.2.2 Health sector inequity in low-income countries

Inequity analysis in health care utilisation can be viewed differently in developing than in developed countries because of the differences in their mix in finance and provision of health care. On the supply side, inequality in access in developing countries can occur because of the uneven distribution of health care providers across regions as well as the discriminatory behaviour against minority groups by health care providers. For instance, Garimoi (2009) points out a better health condition of urban relative to rural population even within a similar level of household incomes, indicating that uneven distribution of health care providers between urban and rural areas can reinforce rural-urban inequity in health outcomes. Discriminatory behaviour against minority groups by health care providers may also influence inequity in health care utilisation even if the provision of health care services is fairly distributed (Dommaraju et al., 2008; Babu et al. 2001). On the demand side, differences in ability to pay undermine health care utilisation by poor and minority groups. Leon and Walt (2001) argue that unequal access to health care between poor and rich is a contributory factor in producing health inequality in low-income countries.

Another contributing factor of health sector inequity in low-income countries different from that in developed countries is in the mix of the financing sources. In developed countries, health care is generally financed from four sources: taxes,

social insurance, private insurance and OOP payments (Wagstaff and Doorslaer, 2001). The contribution of each source to the total health care finance varies across countries. In developed countries, the majority of the population are employed in the formal sector and their contribution via general taxation or social insurance contribution, as well as via private insurance, will cover any future treatments. As a result, they are less likely to face OOP payments when they are in need of health care services. In developing countries, on the other hand, most of the population relies on subsistence agricultural work for their livelihood and are less likely to afford private health insurance. Private health insurance is basically negligible. Even in a rapidly growing economy like India more than 95% of the population does not have private health insurance (Ghosh, 2011). Additionally, low levels of tax revenue are a constraint to the governments of these countries for allocating part of their budget to the provision of health care services. For instance, the contribution of the public sector on total health care expenditure in some sub-Saharan countries is lower than 25% and the coverage of private insurance is limited (WHO, 2009). According to WHO (2009), OOP payments stand as the single largest source of health care financing in low-income countries and thus is an important factor to take into account in the equity analysis of health care utilisation.

4.2.3 Wealth, health care utilisation and health outcomes

A low level of health care utilization by minority groups and deprived communities, which can partly be attributed to their lower level of wealth, can lead to a relatively inferior level of health outcomes in them. Banerjee et al. (2004) investigate the correlation between wealth and health (self-reported health status as well as some direct measures of health such as body mass index hemoglobin level etc.) employing data from Indian village, Rajasthan. Their finding indicated that individuals from

the lower third per capita income have, on average, lower levels self-reported health status as well as other direct measures of health than those in the upper third per capita income. This study also shows the quality of health care services to appear positively correlated with individual health after controlling for access factor (distance), wealth and demographic characteristics (age gender). An international comparison of socioeconomic status and health supports the conclusion derived in Banerjee et al. (2004). For instance, Anne and Deaton (2005) compare similar health measures between people from an Indian village, Rajasthan and the African city, Khayelitsha, where the latter being relatively well in terms of socioeconomic status. Their results show African to have better height and body mass index compared to Indians. However, they did not find subjective measures of health to be better among African relative to Indians. Authors acknowledge this as a contradicting result and argue that there are tendencies of better off people to report inferior health than the poor people especially in developing countries. Positive correlationship between health and wealth suggests that an economic growth alone may not lead to the expected level of improvement in public health of a country in the absence of other factor such as education and institutional improvements (Deaton 2006). Therefore, not only the growth but an equitable distribution of income enhances public health. Comparing income inequality and health outcomes across 22 developed countries, The Equality Trust (an organization based in the UK) finds a higher infant mortality rate and lower life expectancy in countries with more unequal income distributions (The Equality Trust, 2013).

Deaton (2011) argues that childhood inequalities in social status such as parental education and income are the key factors associated with adulthood health. Therefore, public intervention that lessens the inequalities in parental circumstances may lead to the reduction on health inequalities. There are several empirical studies that

show deprived communities and minority ethnic groups are likely to have poorer health outcomes such as life expectancy, child mortality rate and self-assessed health status relative to the wealthiest or dominant group. For instance, African Americans are found to have poorer health and higher mortality rates relative to whites in the United States (Hummer, 1996; Kriger, 1987). In the United Kingdom, migrants of South Asian descent have an ischaemic heart disease-related mortality rate that is 1.5 times higher than that of the rest of the population (Charturbedi, 2003). Mortality rate is higher among Scheduled Castes (SCS) and Scheduled Tribes (STS) relative to the remaining population in India (see Baru et al., 2013; Subramanian et. al., 2006).[3] As in India, the privileged *Tagadhari* caste in Nepal have shown better health indicators, measured as life expectancy and child mortality rates, relative to those discriminated against groups such as Dalit and indigenous people (Bennet, 2005).[4]

4.2.4 Caste and health outcomes

An important but as yet less explored aspect of inequity analysis across social groups is caste inequity in health. Although the caste system has been abolished in many countries, including Nepal, it is believed to perpetuate social inequality in countries with a historically caste-based society through the intersection of different socioeconomic spheres such as economic, cultural, social and political.[5] A caste system has embedded a rigid social hierarchy that is stronger than social classifications such as ethnicity or race (Cox, 1944; Jeffrey, 2001). A caste system imposes a social division of labour that limits low castes to low-paid manual jobs. This is something

[3]Scheduled Castes and Scheduled Tribes are two groups of historically disadvantaged people recognized in the constitution of India.

[4]Dalit and indigenous people represent the *Pani Nachalne* and *Matwali* castes, respectively, in the caste categories of this study.

[5]For instance, the ownership of agricultural land is still dominated by high castes in India (Jaffrey, 2001).

not likely to appear in other forms of social classifications. The direct implication of this is that low castes are indirectly discouraged from obtaining a better education and consequently have a lower income, therefore reinforcing the deficiency in health variables. Low castes may also face direct occupational effects on their health since they will be constrained to manual and more hazardous jobs. In India, lucrative white-collar jobs are in the hands of the high-caste population (Jaffrey, 2001). Figure 4.1 shows that more than 80% of white-collar (professional and clerical) jobs have been occupied by the dominant caste in Nepal.

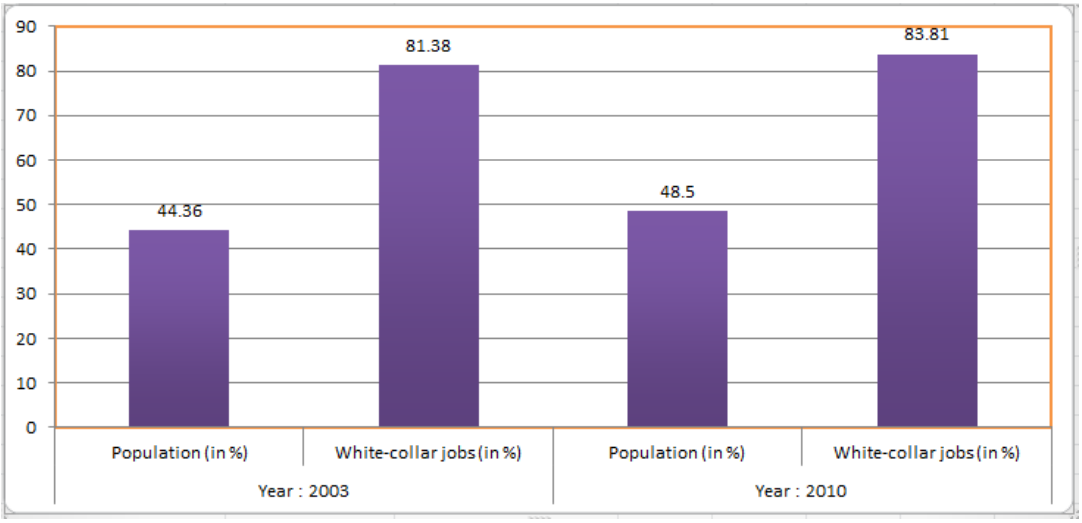

Figure 4.1: White-collar jobs by dominant caste
Source: NLSS 2003 and 2010.

Inequity in health care utilisation across castes becomes acute as OOP stands as the main source of health care financing. As low-caste groups end up with a relatively inferior socioeconomic status they possibly forgo medical treatment or choose poor-quality health care providers because of the impact that OOP expenditure has on poverty. Kawachi et al. (2005) argue that inferior social capital in disadvantaged groups relative to dominant groups contributes to differences in health care utilisation. Additionally, Deaton (2011) argues that childhood inequalities, inequalities in parental circumstances such as income and education, as a key factor to influence

an individual's adulthood health. Therefore, caste-based stratification may depict even a stronger impediment in utilising health care services and thus boosts inequity in health outcomes relative to that arising in ethnic/racial groups.

There are few studies on caste inequities in health carried out in India and Nepal. Roy and Howard (2007) look into the share of OOP payments over household consumption and show that it increases with ability to pay. However, by comparing this relationship with social codes, these authors show that SCS and STS spend less in OOP payments compared to others. Dommaraju et al. (2008) evaluate the effect of caste on child mortality rate and the utilisation of maternal health care services in rural India. They show that low-caste children face a higher risk of death, while low-caste women are less likely to use antenatal health care services compared with the upper castes. Child malnutrition in India is also more common in SCS and STS than in the rest of the population. Lower wealth, education and use of health care services in these groups relative to that in dominant groups are found as major factors that influence child nutritional outcome (Van de Poel and Speybroack, 2009).

To sum up, previous studies show that inequity in health care utilisation exists in all regions irrespective of their level of socioeconomic development. More importantly, attention has not yet been given to caste inequity in health in Nepal, a country that has experienced an age-old legally imposed caste-based division of labour. Nepal, being one of the poorest countries in the world and one where health care expenditure predominantly depended on OOP payments, is striving for health care reform that promotes equity. Furthermore, none of the previous studies has examined, at least to my knowledge, the determinants of health outcomes across castes to see whether disadvantaged castes end up with inferior health outcomes compared to dominant castes.

4.3 Health care system in Nepal

Nepal has a commitment of providing primary health care (PHC) for all segments of the population (Subedi, 1992). Nepal's first long-term health plan (LTHP, 1975-1990) was designed to ensure access to PHC for the entire population, particularly targeting the rural population who had limited access to modern urban health care facilities. LTHP had several strategic approaches - development of infrastructure, community participation, multi-sectoral coordination, mobilizing resources, and decentralization of planning and management of health care - all in order to achieve universal coverage of PHC services (WHO, 2007). The second long-term health plan (II-LTHP, 1997-2017) further focused on socially marginalized and underprivileged groups of people such as women, children, the rural population and the poor. The II-LTHP explicitly defined 20 essential health care services (EHCS) to be provided for the effective delivery of PHC to these target groups (ibid).

Nepal's recent periodic development plans set the priority of the sectoral programs and provided a base for allocating funds across ministries from the national budget and integrating the health policies described above. For instance, the ninth plan (1996-2001) set poverty eradication as its main development objective and introduced target programs to reduce poverty. This periodic plan vowed to improve equity while implementing health-related programs. In particular, economically and socially disadvantaged castes and communities were described as the target groups in implementing the health sector programs. This plan employs an integrated approach for health care services to be provided by district hospitals, primary health care centres, health posts and sub-health posts. Health posts and sub-health posts are the lowest level of health care providers and are extensively networked throughout the country. These provide primary care and referral services at village level.

The objective of poverty alleviation was further emphasized in the tenth plan

(2002-2007) by declaring poverty eradication as the sole development objective.[6] In particular, it acknowledges gender and caste/ethnic disparities in poverty and again includes them in the target group.

There have been several policy interventions within the tenth periodic plan to achieve EHCS goals effectively. One of these was the introduction of the Nepal Health Sector Program-Implementation Plan (NHSP-IP) 2004-2009. NHSP-IP has two features. First, it aims to improve coverage and increase the quality of EHCS while at the same time promoting access. Second, additional actions are laid over the EHCS program by providing financial support to the poor and vulnerable segment of the population (Ministry of Health/Nepal, 2006). In particular, NHSP-IP imposes new actions over EHCS such as the identification of the poor, expansion of EHCS, subsidized drugs and services and the introduction of social safety net programs. As a result, for instance, free treatment was provided to the poor to tackle diseases such as leprosy, TB, HIV/AIDS and malaria.

In addition, user fees for primary health care services provided by health posts and sub-health posts were abolished in 2007 throughout the country for the poor (those falling under the income poverty line) and for the entire population in the 22 districts with the lowest human development indicator (HDI). It was extended to primary health care centres in 2009. The poor were also offered outpatient, inpatient, emergency services and essential medicines at district-level hospitals free of cost (Ministerial Leadership Initiatives, 2010).

Despite all these provisions, health care services in Nepal are still paid for by users

[6]The tenth plan was itself the replication of country's strategic plans in alleviating poverty, which is also known as the Poverty Reduction Strategy Paper (PRSP). The implementation of target groups programs were more explicitly explained in this plan document. This had four pillars of development strategies to achieve the objective of poverty alleviation: (a) pro-poor economic growth, (b) equitable access to social and economic infrastructures for poor and marginalized groups, (c) social inclusion and target programs and (d) improved governance. Pillars (b) and (c) clearly indicate the national initiative to empower socially and economically disadvantaged groups.

both in the public and private sector, although the cost of utilisation significantly varies between them. The private sector is more expensive relative to the public sector; therefore, the poor and disadvantaged are less able to utilize private health care services and rely mainly on care provided by the public sector (WHO, 2007).

4.3.1 Health care providers

Nepal had adopted a system of public-private partnership in providing health care services. The private sector provides health care services through hospitals, nursing homes and clinics. However, these are mainly focused in the capital (almost half of them are based in Kathmandu), regional headquarters and other urban areas. Figure 4.2 below shows that more than 50% of private health care providers are

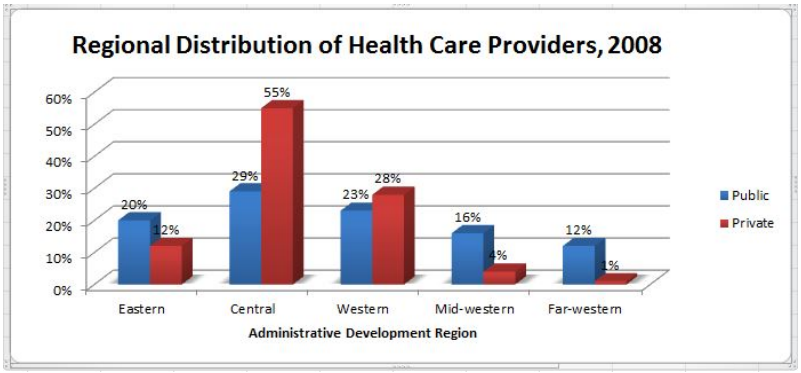

Figure 4.2: Regional distribution of health care providers

situated only in the central region, while only 1% are based in the far western region. Public sector health care providers are almost proportionately distributed across all regions. Non-governmental organizations (NGOs) also provide health care services, albeit on a negligible scale. If economically disadvantaged people are less able to afford private health care services, given that private health care providers charge higher rates, equity in the utilisation of health care services predominantly relies on provision by the public sector to make sure services are provided across income groups and also across regions (urban versus rural presence of providers).

4.3.2 Health Care Finance

The health care sector in Nepal is financed by the government and households OOP expenses (Hotchkiss et al., 1998). However, OOP expenditure is predominant in the share of health care expenditure. In 2002/3, OOP expenses accounted for 62.5% of total health care expenditure, whereas governmental contribution was 16.8%. The remaining health care expenditure was borne by External Development Partners (EDPs) (WHO, 2007). The share of OOP expenses has marginally decreased over time. It accounted for 60.5% of total health care expenditure in 2009 relative to 62.5% in 2003 (Ministry of Health and Population, 2009).[7] The OOP expenses come from relatively well-off people. The richest 5% spent 7.2% of their household budget, whereas the poorest 5% spent only 2.6% (WHO, 2007).

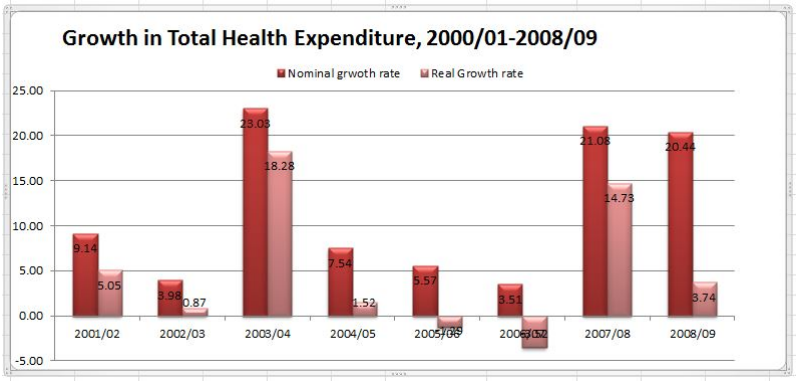

Figure 4.3: Total health care expenditure growth rate (2000-2009).

Total health expenditure (THE) is increasing over time. Figure 4.3 shows that the highest growth rate was achieved in 2003/04, while the lowest was in 2006/07. On average, the THE growth rate over the period of 2000/01 to 2008/09 was 11.8%. However, the average growth rate in real terms for the same period accounts for less than half of the nominal rate.

Share of government spending on total health care finance was in increasing

[7] A comparison of OOP expenses on a yearly basis is not possible since OOP expenses can only be extracted from the National Household Surveys conducted in 2003 and 2010.

trend until financial year 2003/04 but has been decreased in the consecutive years. Eventually there is approximately 5% reduction in the government contribution on THE over the period of 2000/01-2008/09 (decreased to 23% in 2008/09 from 27.5% in 2000/01). Total health care expenditure in relation to the gross domestic products (GDP) remains almost constant throughout the period accounting approximately 5% of total GDP (see figures 4.4 and 4.5).

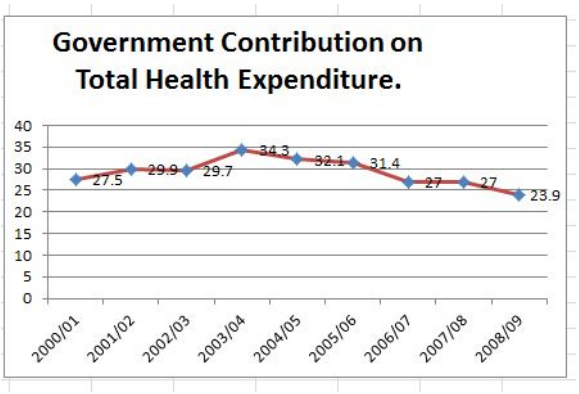

Figure 4.4: Public sector contribution on total health expenditure

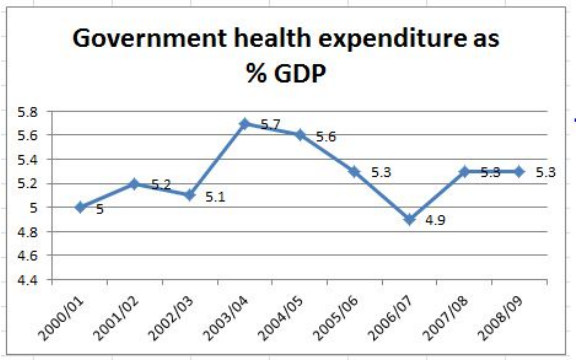

Figure 4.5: Total health care expenditure as % of GDP

4.4 Methodology

4.4.1 Inequity in health care utilisation

This study, first tests for the presence of inequity in health care utilisation across castes in Nepal. It examines differences in health care utilisation, proxied by OOP expenditure arising for those individuals that have the same needs. However, many individuals did not have OOP expenses as they did not suffer from any illness and thus it is imputed as zero. This imputation strategy led to a sample with many zeros in dependent variable. A tobit regression model that accounts for many zeros in predicting expected mean value of OOP expenses is used for this reason. This

model allows non-respondents of OOP expenses to remain in the sample and avoids the issue of sample selection bias. Defining caste categories $j = t, m, p$ (t=*Tagadhari*, m= *Matwali* and p=*Pani Nachalne*) a tobit regression model (Wooldridge, 2002) can be specified as,

$$hc_i = \begin{cases} hc_i^* & \text{if } hc_i^* > 0 \\ 0 & \text{if } hc_i^* \leq 0 \end{cases} \quad (4.1)$$

where hc_i^* is the latent variable: $hc_i^* = \alpha + \beta h_i + \delta X_i + \eta C_{ji} + u_i;\ u_i \sim N(0, 1)$

Need for health care utilisation is defined by binary variables that capture if individuals had any chronic illness, non-chronic illness and were not ill. These variables take value one if an individual falls in each categories and zero otherwise. In addition, two other variables are constructed to capture the effect of severity of illness for those who reported as facing chronic illness. The first variable represents the years of chronic illness started which is divided into four categories while the second is the work-missed days due to the chronic illness. The latter variable is classified into three categories.[8]

X_i contains a set of explanatory variables including socioeconomic (income and education), demographic (age, sex, marital status), geographical (rural/urban and district dummies) and household size variables associated with an individual i. u_i represents the normally distributed error term associated with individual i. The definition of the variables is described in Appendix 3.A.[9]

Since the government abolished user fees in some districts and the utilisation of health care in this analysis is proxied by OOP payments, the model above is likely

[8]Both the year of illness started and work-missed variables are available only for those who face chronic illness.

[9]Two proxies are used for income variable: per capita food and frequently purchased non-food expenses and aggregate household data for the price of land holding.

to produce biased results.[10] Results could be misleading if there is a significant difference among castes in utilising such free services. Information on access to free care is not available. To correct for this potential bias, a binary variable, *Lhdi*, has been included as an additional explanatory variable in the regression model. *Lhdi* has a value of one if an individual belongs to one of the 22 districts of Nepal associated with the lowest HDI and zero otherwise.

Following Wagstaff and Doorsalaer (2002) caste-inequity in health care utilisation can be tested by estimating the latent variable model defined in equation 4.1 for each caste group as,

$$hc_i^* = \begin{cases} \alpha_t + \beta_t h_i + \delta_t X_i + u_i & \text{if caste is } t, \\ \alpha_m + \beta_m h_i + \delta_m X_i + u_i & \text{if caste is } m, \\ \alpha_p + \beta_p h_i + \delta_p X_i + u_i & \text{if caste is } p, \end{cases} \tag{4.2}$$

where α is a constant term and β is the coefficients for binary variables representing need for health care utilisation (Other illness as reference category).[11] If individuals with similar needs were treated alike irrespective of their caste association the sum of the two coefficients will be equal to each categories; $\alpha_t + \beta_t = \alpha_m + \beta_m = \alpha_p + \beta_p$. Alternatively, a single equation model with interaction between caste and need variables can be estimated as,

$$hc_i^* = \pi_0 + \pi_1 C_{ji} + \pi_2 h_i + \pi_3 C_{ji} h_i + \pi_4 X_i + u_i \tag{4.3}$$

where, C_{ji} is caste binary variables (*Tagadhari*, t as reference group) and null hy-

[10]The three-year interim plan (2007/08-2009/10) announced the abolishment of user fees for health care services provided by sub-health posts, health posts and district-level health care centres in the 22 districts with low HDI. The list of districts with different HDI levels is presented in Appendix 1.B.

[11]Not ill variable could be a suitable reference category since chronic illness and other illness both represents need for health care utilisation. However, this was not possible since this variable has all zeros in dependent variables.

pothesis of $\pi_1 = 0$ and $\pi_3 = 0$ can be tested for the presence of caste inequity in health care utilisation.

4.4.2 Caste differential on health outcomes

Following the test of caste inequity in health care utilisation, this study extends the estimation strategy to evaluate the determinants of caste differential in health outcomes. In particular, this part looks into whether low castes end up having lower health outcomes relative to the dominant caste. Self-assessed health (SAH) is used as a measure of health outcome and it is dependent variable. SAH has a category s taking value 1, 2, 3 and 4 respectively for *Very poor, Poor, Fair* and *Good* in 2003 and *Poor, Fair, Good* and *Excellent* in 2010.

The definition in SAH changed between 2003 and 2010. In 2003 there were four possible alternative responses: *Very Poor, Poor, Fair* and *Good* to assess own health status. However, in the 2010 wave SAH was categorized as *Poor, Fair, Good* and *Excellent*. This introduces a constraint to make a comparison and draw any conclusion on how perceiving health across castes might have changed over time. However, the change of wording in SAH categories only restricts to analyse the dynamics of caste-associated SAH but still allows to evaluate the impact of caste on health outcomes. Therefore, it is not considered as a crucial limitation, at least for the purpose of this study.

Consider an ordered probit model (OPM) (Wooldridge, 2002) for the ordered responses s using a latent variable model specification.

$$y_i^* = \gamma X_i + \lambda C_j + e_i,\ e_i \sim N(0,1), \tag{4.4}$$

where y_i^* is a latent variable representing true health of an individual i that belongs to caste $j = t, m, p$, X_i is the set of explanatory variables associated with individual

i, γ is the vector of estimated parameters[12], C_i are dummy variables for identifying caste, λ represents differences in health status among castes and e_i is the error term associated with individual i.

Consider cut points (threshold parameters) $\mu_1 < \mu_2 < \mu_3$ such that

$y = 1$, if $y^* \leq \mu_1$,

$y = 2$, if $\mu_1 < y^* \leq \mu_2$,

$y = 3$, if $\mu_2 < y^* \leq \mu_3$,

$y = 4$, if $y^* > \mu_3$.

The conditional distribution of y^* given X and C can be computed as

$$
\begin{aligned}
P(y=1|X,C) &= P(y^* \leq \mu_1|X,C) = P(\gamma X + \lambda C + e \leq \mu_1|X,C) = \Phi(\mu_1 - \gamma X - \lambda C), \\
P(y=2|X,C) &= P(\mu_1 < y^* \leq \mu_2|X,C) = \Phi(\mu_2 - \gamma X - \lambda C) - \Phi(\mu_1 - \gamma X - \lambda C), \\
P(y=3|X,C) &= P(\mu_2 < y^* \leq \mu_3|X,C) = \Phi(\mu_3 - \gamma X - \lambda C) - \Phi(\mu_2 - \gamma X - \lambda C), \\
P(y=4|X,C) &= P(y^* > \mu_3|X,C) = 1 - \Phi(\mu_3 - \gamma X - \lambda C)
\end{aligned}
$$

where $\Phi(.)$ is the normal cumulative distribution function.

Equation 4.4 assumes identical cut points for all individuals. Which implies that estimated coefficients for response variables do not varry across sub-samples. However, recent literature suggests that individuals heterogeneity in perceiving own health status can affect thresholds and thus the traditional OPM might produce spurious results (Greene et al., 2008; Pudney and Shields, 2000). In fact a test for the parallel regression line assumption rejects the null hypothesis of equality of coefficients across response categories. Thus a generalized ordered probit model (GOPM) has been estimated in order to avoid possible bias in OPM.

[12]Omitted categories for binary variables are identically used as those in the regression model.

GOPM allows cut points to depend on covariates,

$$\mu_{is} = \tilde{\mu}_s + \psi_s X_i + \xi_s C_i, \tag{4.5}$$

where ψ_s and ξ_s represents the influence of covariates on cut points and $\tilde{\mu}$ represents cut points under the assumption of parallell regression line (OPM). Following equation (4.5) the cumulative probability of GOPM can be written as

$$Prob(y \leq s|X) = \Phi(\tilde{\mu}_s + (\psi_s - \beta)X + (\xi_s - \lambda)C). \tag{4.6}$$

This approach estimates 3 binary probit models. The first model (Model 1) estimates SAH category 1 vs. SAH categories 2, 3 and 4; the second model (Model 2) estimates categories 1 and 2 vs. categories 3 and 4, and finally the last model (Model 3) estimates categories 1, 2 and 3 vs. category 4.

4.5 Data and variable specification

4.5.1 Data

This study employs two waves of the National Living Standards Survey (NLSS) data from Nepal carried out by the Central Bureau of Statistics of Nepal (CBS/N) in technical support of the World Bank. The first survey, carried out in 2003, collected information from 5240 households (28110 individuals). The second survey, carried out in 2010, provides information for 5998 households (28670 individuals). Both surveys include a wide array of economic, demographic and health-related information both at the household and the individual level. With particular reference to health-related information, both surveys incorporate information such as whether individuals suffer from any chronic illness, and whether household members have

faced any diseases, injury or illness in the past month. Information is also provided on whether the individuals who had an illness consulted health care providers and, if so, the type of health care providers consulted. Cost incurred in receiving care in current year is reported by cost components such as consultation fees, medicine cost, travel cost and other expenses in both surveys. This study uses subsamples of 19490 and 20979 individuals in 2003 and 2010, respectively. There is a slight drop in sample size used in this analysis relative to the full sample because of missing information on SAH categories. It is mainly because of some individuals were not presented at the time of survey and it was not possible for respondents to state other individual's SAH status.

Descriptive statistics for all variables used in the analysis are presented in Table 4.1 for 2003 and Table 4.2 for 2010. Both tables show a *prima-facie* evidence that the dominant caste, *Tagadhari*, enjoys a relatively better socioeconomic position as measured in log of income, price of landholding and has a higher level of educational attainment than the other groups.[13] The *Pani Nachalne* group, which is considered to be the most marginalized group and was also considered untouchable (described as "polluting" people in terms of ritual purity under the Hindu system), displays the lowest level of socioeconomic status. The *Matwali* is situated on the intermediate level.

The *Tagadhari* appeared to spend a considerably higher amount on health care services than the other two groups in both periods. This group spends 9.1% (.832-.741) more than the national average and 29.7% (.832-.535) more than the amount spent by the *Matwali* group, which exhibited the lowest average amount of health care expenditure in 2003.

OOP differentials between the *Tagadhari* and *Matwali* is further widened in

[13]Income is proxied by the sum of per capita food expenses and frequently purchased non-food expenses.

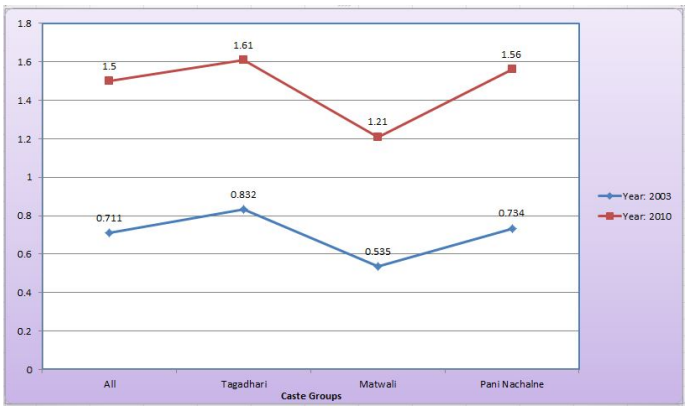

Figure 4.6: OOP Expenses by castes

2010 accounting 40% (1.61-1.21) relative to 29.7% in 2003. However, there is a considerable contraction in the gaps of OOP expenses between the *Tagadhari* and *Pani Nachalne* groups in this period. (see figure 4.6).[14]

In 2003, 59.2% of respondents reported their SAH status as *Good*, followed by 39% reporting *Fair* and 1.7% reporting *Poor* SAH. (SAH is categorized as *Very poor, Poor, Fair* and *Good* in 2003). The dominant group reports present health status as *Good*, slightly higher than the national average as well as than the other groups. The *Pani Nachalne* group has the lowest proportion of individuals reporting *Good* health.

In contrast to 2003, the *Tagadhari* group reported *Excellent* SAH lower than other caste groups in 2010. The proportion of reporting *Excellent* SAH in this period is the highest for the *Matwali* group. The *Matwali* group displays 64.8%, 33.5% and 1.6% in *Excellent, Good* and *Fair* Sah categories, respectively. Same measures accont for 55.4%, 42.1% and 2.4% for the *Pani Nachalne* and 53.9%, 43.9% and 2.1% in case of *Tagadhari* group. A negiligble proportion of individuals have reported their SAH status being in the lowest categories (*Very poor* in 2003 and *Poor* in 2010) in both periods. Figure 4.7 shows a caste comparison of SAH for

[14]These are differences in the log of OOP expenses and thus can be interpreted as the percentage difference.

both periods.

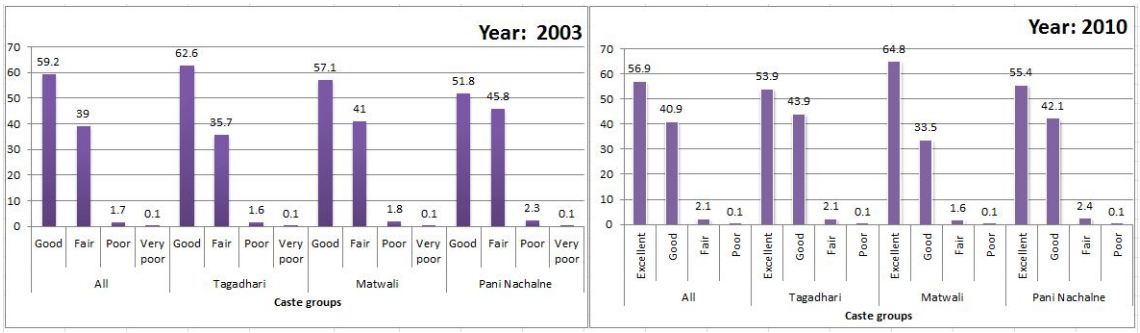

Figure 4.7: Comparison of health status by castes

The proportion of individuals with illnesses, which is proxied for need for health care utilization in this model, varies across castes. The *Tagadhari* group reports a higher proportion of having chronic illness in both periods relative to other two caste groups in comparison. For instance, 7.1% and 13.3% individuals from this caste group has reported to have chronic illness in 2003 and 2010, respectively. These measures stand at 4.8% and 10% for the *Matwali* and 5.7% and 10.6% in case of the *Pani Nachalne* group. On the other hand, the proportion of individuals reporting other illness (non-chronic) is higher for both comparison groups than the *Tagadhari* group in 2010. This measure is almost similar across casts in 2003.

The duration of chronic illness and work missed due to this vary across castes. The *Tagadhari* group reports to face chronic illness for longer period than other caste groups in both periods. Similarly, the *Tagadhari* also reports more work days missed due to chronic illness relative to other castes. In 2003, 96.9%, 1.4% and 1.7% individuals from the *Tagadhari* group reported missing work by one week, one month and more than one month, respectively. The same measure stands at 97.8%, 1.1% and 1.1% for the *Matwali* and 96.9%, 1.8% and 1.3% for the *Pani Nachalne* caste. A similar pattern of work missed due to chronic illness across caste groups can be observed in 2010.

4.5.2 Variable specification

This study estimates a tobit model for assessing caste inequity in health care utilization and ordered probit as well as generalised ordered probit models to evaluate the impact of caste on health outcomes. The first model uses the log of OOP expenses as the depended variable representing the sum of consultation, diagnosis, medicines and travel cost. A direct measure of utilization such as physician visit, number of hospital visit or use of outpatient or ambulatory care services could represent a more appropriate measure for health care utilisation compared to the use of OOP expenses as proxy for these measures. However, this information is not available in data set used in this analysis. The second model uses subjective measures of health status (SAH) as the dependent variable. SAH is extensively used in the field of public health research and empirical evidence have shown that SAH and objective health status are positively correlated (Wu et al., 2013). In this analysis SAH is categorised as *Very poor, Poor, Fair* and *Good* in 2003 while there is a slight change in the wording of SAH categories in 2010, as mentioned earlier.

The covariates are presented in Appendix 3.A. The selections of covariates are similar to those already used in empirical works. The categorical age variables are used to capture the age effect on the occurrences of illnesses. These samples represent individuals with all age categories. Age category less than 34 years is used as the reference group since very few individuals in the lower age group, particularly within the age of 20 years, have reported to face chronic illness which is an important variable to influence both dependent variables used in this analysis.[15] Two variables are used as the proxy for household income. Lincome is the log of household gross income proxied by the sum of food expenses and frequently purchased non-food expenditure by households. Since only household consumption may not reflect their

[15]Reference age category includes 20.27% and 17% of individuals facing chronic illness in 2003 and 2010, respectively.

actual wealth status an additional variable is used to control for income variation across households. The second proxy for household income is the price of land holdings by each household. These variables are crucial in explaining the OOP expenses as well as health status since wealthier people may not only have higher ability or willingness to pay for health care services but they may also have better knowledge about health care. Similarly, as the number of individual in a household lessens the resources available per person a continuous variable to represent household size, Hhsize, has also been included as an additional covariate.

Rural and district dummies are included to capture the possible behavioural differences in individuals living in different districts and rural/urban areas. These variables may also capture the rural/urban and geographical heterogeneity in the supply of health care services. Gender and marital status are also included since both variables can influence health.

Education is used as an additional covariate because it may not only influence efficiency in the production of health but also the propensity of seeking care. Education variables are computed from the highest level of schooling completed by an individual. However, those who did not reported the highest level of schooling completed but responded as they never attended school in the past in the following survey questionnaire their educational level has been imputed as zero. Four categories of the level of education are used: illiterate (individuals without formal education), primary (individuals having 1-5 years of schooling), high school (individuals with 6-10 years of schooling) and university (individuals with more than ten years of schooling).

Finally, other covariates in this study include three types of binary variables to capture the need for health care utilisation. The first variable, chronic illness variable, carries value of 1 if an individuals has reported having chronic illness and

0, otherwise. Other illness variables contains value 1 if an individual reports having illness other than chronic illness and 0, otherwise. Not ill variables has value 1 if an individual reports to have neither types of illness and 0, otherwise. It has also been acknowledged that the severity of illness can have influences on health care utilization as well as on health outcomes. Therefore, two types of additional variables are used to capture these effects. The first represents the duration that an individual is suffering from chronic illness which is divided into four categories: suffered from one or less year, less than five years, less than ten years and more than ten years. The second is the work- days missed due to the chronic illness. This variable represents three categories: work missed less than a week, more than a week but less than a month and work missed more than a month.

4.6 Empirical results

4.6.1 Equity analysis

4.6.1.1 Tobit regression results

This study uses censored regression model to analyse inequity in health care across castes. Regression results are listed in Table 4.3 for 2003 and in Table 4.4 for 2010. Results from the *Tagadhari*, *Matwali* and *Pani Nachalne* sub-samples are reported in columns 1, 2 and 3, respectively. Columns 4 and 5 in these tables represent results from the pooled sample with caste dummies and the interaction between caste dummies and need variables as additional covariates, respectively. Coefficients for district dummies are not presented in order to save space.

In 2003, results show that log of income, chronic illness and the severity of illness (captured by work missed days due to illness) have significant positive impact on health care utilisation by each caste groups although the magnitudes of each variable

vary across castes. Log of income coefficient is considerably high in the *Matwali* sub-sample implying that this caste has higher income elasticity to health care utilisation than the others. Chronic illness coefficient is the highest for the *Tagadhari* group. The *Pani Nachane* group shows the lowest and statistically insignificant coefficient for chronic illness in this period. Rural and married coefficients are statistically significant only in *Tagadhari* sub-sample while *Lhdi* coefficient is significant only in the *Matwali* sub-sample. The remaining control variables generally show statistically insignificant results. Results from pooled sample (column 4) show that both castes, the *Matwali* and *Pani Nachalne*, have a negative and statistically significant coefficient. In the fifth column, both castes show negative coefficients for interaction term (caste dummies and need variables) indicating that low castes spend less relative to the dominant caste for the treatment of chronic illness relative to other illness. However, this coefficient is not statistically significant for *Matwali*. In the same column caste variables show negative coefficients but only the the *Matwali* coefficient is statistically significant. Therefore, in line with the hypothesis outlined in section 4.4 results indicate a presence of caste inequity in health care utilisation in Nepal which favours the dominant group.

Results for 2010 are presented in Table 4.4. There is a significant drop in the impact of income on health care utilisation. These coefficients other than for *Matwali* are statistically significant. Rural and married variables depict a similar trend in influencing health care utilisation as in 2003. Additionally, household size variable became statistically significant in this period which shows a positive impact on health care utilisation. *Lhdi* coefficients other than in pooled sample are statistically insignificant in this period.

As in 2003, the severity of illness variables (captured by work missed days due to illness) show significant positive impact on health care utilisation in this period.

Additionally, other proxy for severity of illness (years of illness started) also shows a statistically significant coefficient for the *Tagadhari* and *Matwali* sub-samples in this period. It shows that the longer the duration of illness the lower the usilisation of health care services. It is not surprising since recent illness requires various diagnostic and consultation cost along with other costs such as medicine and travel expenses which may not require in the same level for the older illness. Chronic illness coefficients considerably increased in this period relative to 2003. In contrast to 2003, the *Matwali* caste shows the highest coefficients for chronic illness. In the fourth column, caste variables show similar coefficients as in 2003 implying that both castes in comparison utilise less health care services compared to the dominant caste. In the fifth column, results show that interaction coefficients for both castes are negative but this coefficient for the *Matwali* caste is not statistically significant. In the same column, both castes show negative coefficients. However, the *Pani Nachalne* is not statistically significant. The results for 2010 show that both castes the *Matwali* and *Pani Nachalne*, seem to fare inequity in health care utilisation compared to the dominant caste. Although not comparable, the magnitudes of caste-inequity in health care utilisation between two periods indicates that it is decreasing over time.

4.6.1.2 Marginal effects

This subsection analyses marginal effects to see how each variables are likely to influence health care utilisation. As this study aims to evaluate caste differentials in OOP expenses for health care services marginal effects conditional on being uncensored (those with positive OOP expenses) are reported in Table 4.5 for both periods. Column 1 in this table reports marginal effects without interacting caste and need variables followed by results from interaction model in column 3 for 2003. Same

results for 2010 are presented in columns 3 and 4.

Results show that an increase in the log of income by 1% is likely to increase OOP expenses approximately by 2% in 2003. The impact of log of income on OOP expenses is dropped approximately to half in the latter period. These coefficients are significant at 1% level. Caste variables show negative coefficients in both periods indicating that low castes, the *Matwali* and *Pani Nachalne*, are likely to spend less relative to the dominant caste, *Tagadhari*. While interaction terms are included as additional control variables, the *Matwali* coefficient slightly drops in 2003 but increases in the latter period while the *Pani Nachalne* coefficients become statistically insignificant in both periods. Interaction term (interaction with chronic illness) coefficients are negative and statistically significant for the *Pani Nachalne* but not for *Matwali*. It implies that the lowest caste, *Pani Nachalne*, spends less in face of illness whereas the intermediate caste, *Matwali*, spends less as they face less illness relative to the *Tagadhari* group.

As expected, *Lhdi* variable showed negative coefficients in both periods implying that residing in the districts associated with low levels of HDI are likely to spend less in health care utilisation relative to those who live in districts with better HDI. These coefficients are significant at 5% level in both periods. In the latter period, *Lhdi* ciefficients become positive. It was not expected since government announced to abolish users fees in these districts. One could expect the *Lhdi* coefficient to increase in the latter period as government announced to provide health care services free of cost in these districts. It indicates that the government policy of abolishment of user fees has not been effective.

4.6.2 The impact of caste on SAH

In this section I proceed to examine whether there exists a caste influence on health outcomes. The objective of this additional estimation strategy is to examine whether historically discriminated-against groups end up in a relatively inferior level of SAH relative to the dominant group. The ratio of OOP health care expenditure over income instead of the log of income is used in these models (referred to as OOP ratio). All other explanatory variables are the same as before. Results from the OPM are listed in Tables 4.6 and 4.7 for 2003 and 2010, respectively. Each column represents the marginal effect on the probability of reporting s, where s = *Very poor, Poor, Fair* and *Good* for columns 1, 2, 3, and 4, respectively.

Results in Table 4.6 for 2003 show that OOP ratios are negatively correlated with *Good* but positively correlated with *Fair* health status. However, these coefficients are statistically insignificant. Rural inhabitants and individuals from larger households are likely to report their health status as *Fair* and less likely to report as *Good*. These results were expected since rural population may have less access to health care services. In addition, both of these variables can minimise the utilisation of health care services via income effect which might led to the inferior health status with them. Male and married people are likely to report better health status compared to female and unmarried people. No significant difference between the residence of low HDI districts and others is found in terms of reporting health status.

As expected, chronically ill individuals are likely to report their SAH as *Very poor, Poor* and *Fair* and less likely to report as *Good*. On the other hand, those who were ever ill report better health status. All coefficients for age categorical variables show a positive sign for the lower three SAH categories but a negative sign for SAH status categorised as *Good*. These coefficients are statistically significant

at 1% level. It implies that elders are likely to report inferior health status relative to those younger than 35 years of age (reference category).

Years of illness started variables show expected but statistically insignificant results while work missed variables are as per the expectation and statistically significant. It shows that persons who missed work more than a week are less likely to report their health status as *Good* relative to those who missed work less than a week due to chronic illness. Finally, both cast variables depicted positive and statistically significant coefficients for inferior SAH (*Very poor, Poor* and *Fair*) but negative coefficients for better SAH (*Good*) indicating that historically discriminated-against caste groups are likely to end up in an inferior health outcomes.

In 2010, the effects of OOP ratios on SAH show a similar effect as in 2003. The coefficients for OOP ratio are statistically significant at 1% level in this period. Similarly, education, age, chronic illness, not ill and severity of illness (years of illness started and work missed variables) variables show a similar behaviour as in 2003. The *Pani Nachalne* caste shows negative coefficient for *Excellent* SAH and positive coefficients for lower three SAH categories. These coefficients are significant at 5% level. In contrast, the *Matwali* coefficients are opposite for *Excellent* and negative for all other SAH categories and are statistically significant at 1% level. These results are indicative of a significant decrease in caste inequity in health outcomes over time. However, the lowest caste (*Pani Nachalne*) still lags behind the dominant caste in terms of health outcomes. The intemediate caste, *Matwali*, seems to have a considerable improvement in health outcomes over time superseding even to the dominant caste, *Tagadhari*.

As discussed in section 4.4, OPM is based on proportional odds or the parallel line regression assumption. In other words, OPM assumes that the coefficients that describe the relationship between dependent and response variables are constant

across categories. However, there may be group differences in perceiving own health status across castes. A likelihood ratio test is carried out to evaluate whether the parallel regression assumption holds. The null hypothesis is that of identical coefficients of explanatory variables across categories. The test rejects the null hypothesis, suggesting that the assumption of parallel line regression is violated.[16] Therefore a GOPM, that allows coefficients for explanatory variables to vary across categories, is estimated. Results are presented in Tables 4.8 and 4.9, respectively, for 2003 and 2010. Marginal effects are reported instead of coefficients.

Consider Model 1 for 2003, which estimates a probit in which the dependent variable equals to 1 if individuals report SAH as *Very Poor* and 0 if they report *Poor, Fair* and *Good*. Due to very few respondents reporting their health status as *Very Poor*, many variables have been dropped in this model, resulting in 4747 out of total 19490 observations. Coefficients for explanatory variables other than severity of illness variables generally appear to be statistically insignificant. Severity of illness variables show a positive coefficients indicating that individuals with severe illness have a *Very poor* health status.

Model 2 estimates a probit in which the dependent variables equals to 1 if individuals report SAH as *Very Poor* and *Poor* and 0 if they report *Fair* and *Good*. *Lhdi* and chronic illness variables show positive and statistically significant coefficients. *Lhdi* coefficient is significant at 1% level while the chronic illness coefficnt is significant at 5% level. These results indicate that those living in districts with low HDI and chronically ill are likely to report *Very poor* and *Poor* SAH relative to those living in districts with better HDI and those without chronic illness. Similarly, the coefficient for not ill variable shows an opposite sign to the coefficient for chronically

[16]STATA could not report likelihood ratio test for the identical coefficients for explanatory variables while district dummies were used as covariates possibly due to the larger number of explanatory variables (100 variables including 76 district dummies). Therefore, district dummies are aggregated to six regional dummies as per the administrative division of Nepal.

ill and is significant at 1% level.

Age and education variables also show expected results. These results indicate that elders are more likely to report *Very poor* and *Poor* SAH relative to younger people while educated people are less likely report these SAH compared to those with no education. Years of illness started variables show unexpected but statistically insignificant results. Work missed variables depicted positive and statistically significant coefficients implying that those who missed more days due to chronic illness are likely to report inferior SAH. Both castes coefficients do not appear statistically significant in this model.

Model 3 estimates the probability of reporting SAH as *Very Poor*, *Poor* and *Fair* relative to *Good*. The probit model has the dependent variable a dummy equal to 1 if SAH is *Very Poor*, *Poor* and *Fair* and 0 if SAH is *Good*. Results from this models follow the pattern of Model 2 in terms of the sign of estimated coefficients but a considerable increase in magnitudes. The levels of significance have also been considerably improved in this model. Both caste variables show positive coefficients. However, the *Matwali* coefficient is not statistically significant in this model. This indicates that historically discriminated-against castes, especially the lowest caste, are less likely report their health status as *Good* compared to the dominant caste.

In 2010, the OOP ratio coefficient is statistically significant only in Model 2. As in 2003, many variables are dropped in Model 1 and most of the coefficients are not statistically significant. Coefficients for married dummy, education and *Pani Nachalne* caste variables are statistically significant at 5% level which are consistent, in terms of the sign, with the coefficients obtained in 2003.

In Model 2, OOP ratio, chronic illness, married and not ill dummy variables show statistically significant coefficients (significant at 1% level). The former two variables show positive coefficients which implies that individuals who spend a larger

amount of OOP for health care services and are chronically ill are likely to report inferior SAH relative to those who spend less and are not chronically ill. Coefficients for age variables are positive and statisically significant in general suggesting that elders report an inferior health status compred to those younger than 35 years of age (reference category). Consistent with the 2003 results all education variables show negative and statistically significant results. It implies that educated people are less likely to report inferior health status relative to those with no education. Work missed variables show again statistically significant positive coefficients. Years of illness started variables became statistically significant in this period. It shows that the higher the years of illness started the lesser is likely to report inferior SAH. Both caste variables show a positive but statistically insignificant coefficients in this model.

Model 3 follows similar pattern of results obtained from model 2. The magnitudes of estimated coefficients have considerably increased. In contrast to 2003, the *Matwali* caste shows a negative and statistically significant coefficient. The *Pani Nachalne* however continues to show a positive sign, albeit less in magnitude of the coefficient as well as the level of significance. These results suggest that caste inequity in health outcomes has decreased over time. However, the lowest caste is still not able catch up other castes in terms of health outcomes.

4.7 Conclusion

This study evaluates the inequity in health care utilisation and examines the determinants of SAH across castes in Nepal, which had age-old legally imposed caste-based social divisions in the past. It finds that historically discriminated-against castes utilise considerably fewer health care services compared to the dominant caste in both periods. This indicates low castes face financial impediments to obtaining

health care services. Low castes may opt for lower-quality and cheaper health care services, or they may not complete a full course of medication or treatment due to the financial constraints they face.

An important finding of this study is that a significant portion of caste differentials in OOP expenses for health care services cannot be explained by caste differences in need variables. Even after controlling for illness, caste differentials in health care utilisation remain significant in both periods with some decrease in magnitudes. Since both samples do not include same respondents an explicit comparison of caste inequity between the two periods may not be possible. However, these results are indicative of decrease in caste inequity over time.

Additionally, a positive relationship between household income and the utilisation of health care services indicates a clear link between income inequality and inequity in the utilisation of health care services. For instance, a relative decrease in caste differential in household income in the latter has has led to lower gaps in the utilisation of health care services. Therefore policies that reduce variance in household income positively contribute to the reduction of caste inequity in health care utilisation in Nepal.

This study also finds that historically discriminated-against casts end up in inferior health outcomes. Both castes showed *Very poor* and *Poor* health status relative to the dominant caste in the first period of study. In the second period, caste inequities in health outcomes between the dominant and intermediate castes have been reversed. The lowest caste indicated poor health relative to the dominant caste even in this period albeit there is a significant decrease in its magnitudes.

Finally, whether the government health policy interventions had any effect in reducing caste inequity should be discussed with caution. Nevertheless, this study did not carry out a comparative analysis between the districts with low HDI, where

government abolished user fees, and the rest of the country; it is found that OOP expenses for health care services have relatively increased in those districts with low HDI in the second period. This indicates that inhabitants of those districts might not have used health care services free of cost as announced by the government. Therefore, it can be argued that, though not precisely, the government policy of waiving fees may not have any positive effect on health care utilisation by low castes or by the poor segment of population.

Table 4.1: Descriptive statistics: 2003

Variable	*All*			*Tagadhari*			*Matwali*			*Pani Nachalne*		
	Mean	*Min*	*Max*	*Mean*	*Min*	*Max*	*Mean*	*Min*	*Max*	*Mean*	*Min*	*Max*
Log OOP	.741(.015)	.711	.771	.832(.021)	.790	.874	.535(.025)	.484	.586	.734(.033)	.668	.799
Lincome	9.27(.003)	9.26	9.27	9.39(.004)	9.38	9.40	9.20(.008)	9.18	9.22	8.99(.008)	8.97	9.01
Lnholding('00000)	120(4)	112	129	165(7)	150	181	72(3)	65	78	49(2)	45	53
Rural	.822(.002)	.817	.827	.769(.004)	.761	.777	.918(.003)	.910	.925	.858(.005)	.847	.869
Male	.474(.003)	.467	.481	.474(.004)	.464	.483	.475(.007)	.461	.489	.471(.008)	.455	.487
Married	.428(.003)	.421	.435	.437(.004)	.428	.447	.408(.007)	.394	.422	.425(.008)	.409	.441
Hhsize	6.41(.019)	6.37	6.45	6.21(.024)	6.16	6.26	6.27(.037)	6.20	6.34	7.20(.058)	7.09	7.31
Lhdi	.188(.002).182	.193	6.21(.024)	.205(.003)	.198	.213	.118(.004)	.109	.128	.224(.006)	.211	.238
Age												
0-34	.687(.003)	.680	.693	.670(.004)	.660	.678	.695(.006)	.682	.708	.729(.007)	.715	.743
35-44	.111(.002)	.106	.115	.118(.003)	.111	.123	.106(.004)	.096	.114	.096(.004)	.086	.105
45-54	.085(.001)	.080	.088	.088(.002)	.082	.093	.082(.003)	.074	.089	.079(.004)	.069	.087
55-64	.063(.001)	.059	.066	.065(.002)	.060	.069	.059(.003)	.052	.065	.061(.003)	.052	.068
65-74	.035(.001)	.033	.038	.038(.001)	.034	.042	.039(.002)	.033	.044	.024(.002)	.018	.028
$\geq$ 75	.019(.001)	.017	.020	.021(.001)	.018	.024	.019(.001)	.014	.022	.012(.001)	.008	.015
Educational level												
Illiterate	.465(.003)	.458	.472	.383(.004)	.374	.392	.524(.007)	.509	.537	.634(.007)	.618	.649
Primary	.245(.003)	.239	.251	.230(.004)	.221	.237	.277(.006)	.264	.289	.252(.007)	.238	.266
High school	.187(.002)	.181	.192	.230(.004)	.222	.238	.162(.005)	.151	.172	.091(.004)	.081	.100
University	.103(.002)	.097	.106	.157(.003)	.149	.163	.037(.002)	.031	.042	.023(.002)	.017	.026
Chronic illness-types												
Chronic illness	.062(.002)	.059	.065	.071(.002)	.066	.075	.048(.003)	.041	.053	.057(.003)	.049	.063
Other illness	.112(.002)	.107	.116	.114(.003)	.108	.119	.101(.004)	.092	.109	.120(.005)	.109	.130
Not ill	.826(.002)	.819	.830	.815(.003)	.807	.822	.851(.005)	.840	.861	.823(.006)	.810	.835
Chronic illness started												
Illness started$\leq$ one year	.939(.001)	.935	.942	.931(.002)	.926	.935	.953(.003)	.947	.959	.945(.003)	.937	.952
Illness started$\leq$ five years	.031(.001)	.028	.032	.034(.001)	.030	.037	.022(.002)	.017	.025	.031(.002)	.025	.036
Illness started$\leq$ ten years	.018(.001)	.016	.019	.020(.001)	.017	.022	.013(.001)	.009	.016	.017(.002)	.012	.021
Illness started$>$ ten years	.012(.001)	.011	.014	.015(.001)	.012	.016	.012(.001)	.008	.015	.007(.001)	.004	.010
Work missed												
Work missed$\leq$week	.971(.001)	.968	.973	.969(.001)	.965	.972	.978(.002)	.973	.982	.969(.002)	.963	.974
Work missed$\leq$month	.014(.001)	.012	.015	.014(.001)	.011	.016	.011(.001)	.008	.014	.018(.002)	.013	.022
Work missed$>$month	.015(.001)	.012	.016	.017(.001)	.014	.019	.011(.001)	.007	.013	.013(.001)	.009	.016
Health status												
Good	.592(.003)	.585	.599	.626(.004)	.617	.635	.571(.007)	.556	.585	.518(.008)	.502	.534
Fair	.390(.003)	.382	.396	.357(.004)	.348	.365	.410(.007)	.396	.425	.458(.008)	.442	.474
Poor	.017(.001)	.015	.019	.016(.001)	.013	.017	.018(.001)	.013	.021	.023(.002)	.017	.026
Very poor	.001(.000)	.001	.001	.001(.000)	.001	.001	.001(.000)	.0001	.001	.001(.000)	.000	.002
Obs.		19490			11055			4735			3700	

Notes: Standard errors in parentheses.

Table 4.2: Descriptive statistics: 2010

Variable	*All*			*Tagadhari*			*Matwali*			*Pani Nachalne*		
	Mean	*Min*	*Max*	*Mean*	*Min*	*Max*	*Mean*	*Min*	*Max*	*Mean*	*Min*	*Max*
Log OOP	1.50(.020)	1.46	1.54	1.61(.028)	1.55	1.66	1.21(.036)	1.14	1.29	1.56(.046)	1.47	1.66
Lincome	10.02(.004)	10.01	10.03	10.10(.005)	10.09	10.11	9.95(.008)	9.93	9.96	9.86(.009)	9.84	9.88
Lnholding('00000)	672(36)	600	743	1025(64)	898	1151	268(13)	241	294	146(7)	132	161
Rural	.794(.002)	.789	800	.745(.004)	.737	.753	.901(.004)	.893	.909	.801(.006)	.788	.813
Male	.460(.003)	.453	.466	.462(.004)	.453	.471	.458(.006)	.445	.472	.454(.007)	.439	.470
Married	.710(.003)	.704	.716	.709(.004)	.700	.717	.697(.006)	.685	.710	.731(.007)	.717	.745
Hhsize	5.70(.018)	5.67	5.74	5.44(.022)	5.40	5.48	5.82(.038)	5.75	5.90	6.33(.043)	6.24	6.41
Lhdi	.219(.002)	.213	.225	.223(.003)	.215	.230	.178(.005)	.168	.189	.263(.007)	.249	.276
Age												
0-34	.664(.003)	.658	.671	.642(.004)	.633	.650	.674(.006)	.661	.687	.721(.007)	.706	.734
35-44	.113(.002)	.109	.118	.123(.003)	.117	.129	.108(.004)	.099	.116	.093(.004)	.084	.102
45-54	.091(.001)	.087	.095	.095(.002)	.089	.100	.089(.003)	.081	.096	.084(.004)	.074	.092
55-64	.067(.001)	.063	.070	.070(.002)	.065	.074	.069(.003)	.062	.076	.054(.003)	.046	.060
65-74	.042(.001)	.038	.043	.045(.001)	.040	.048	.039(.002)	.033	.043	.034(.002)	.028	.040
$\geq$ 75	.023(.001)	.019	.023	.025(.001)	.022	.027	.021(.001)	.016	.024	.014(.001)	.010	.017
Educational level												
Illiterate	.362(.003)	.355	.368	.304(.004)	.295	.312	.395(.006)	.382	.408	.492(.007)	.475	.507
Primary	.205(.002)	.199	.210	.174(.003)	.167	.181	.243(.005)	.231	.255	.247(.006)	.233	.260
High school	.217(.002)	.210	.221	.234(.003)	.226	.241	.224(.005)	.211	.234	.153(.005)	.141	.163
University	.216(.002)	.157	.167	.288(.004)	.279	.295	.138(.004)	.128	.147	.108(.004)	.099	.118
Chronic illness-types												
Chronic illness	.120(.002)	.115	.124	.133(.003)	.126	.139	.100(.004)	.091	.106	.106(.004)	.096	.115
Other illness	.157(.002)	.152	.162	.147(.003)	.140	.153	.160(.005)	.150	.170	.184(.006)	.172	.196
Not ill	.723(.003)	.716	.729	.720(.004)	.711	.728	.740(.006)	.728	.751	.710(.007)	.695	.723
Chronic illness started												
Illness started$\leq$ one year	.892(.002)	.888	.896	.879(.002)	.873	.885	.911(.003)	.903	.919	.906(.004)	.897	.915
Illness started$\leq$ five years	.049(.001)	.046	.052	.052(.002)	.048	.056	.043(.002)	.037	.048	.048(.003)	.040	.054
Illness started$\leq$ ten years	.033(.001)	.030	.034	.038(.001)	.033	.040	.024(.002)	.019	.027	.030(.002)	.024	.035
Illness started$>$ ten years	.026(.001)	.023	.028	.031(.001)	.027	.034	.022(.002)	.017	.025	.016(.002)	.012	.020
Work missed												
Work missed$\leq$week	.985(.000)	.983	.986	.986(.001)	.984	.988	.988(.001)	.985	.991	.978(.002)	.973	.982
Work missed$\leq$month	.012(.000)	.010	.013	.012(.000)	.009	.013	.009(.001)	.006	.011	.016(.001)	.011	.019
Work missed$>$month	.003(.000)	.002	.003	.002(.000)	.001	.002	.003(.000)	.001	.004	.006(.001)	.003	.008
Health status												
Excellent	.569(.003)	.562	.575	.539(.004)	.529	.547	.648(.006)	.635	.660	.554(.007)	.538	.569
Good	.409(.003)	.402	.416	.439(.004)	.429	.447	.335(.006)	.321	.347	.421(.007)	.405	.436
Fair	.021(.000)	.018	.022	.021(.001)	.018	.023	.016(.001)	.012	.019	.024(.002)	.019	.029
Poor	.001(.000)	.000	.001	.001(.000)	.000	.001	.001(.000)	.000	.002	.001(.000)	.000	.001
Obs.		20979			11823			5242			3914	

Notes: Standard errors in parentheses.

Table 4.3: Tobit regression: 2003

Dependent variable: OOP expenses					
Variable	*Tagadhari*	*Matwali*	*Pani Nachalne*	All	Interaction
	1	2	3	4	5
Lincome	.998***(.184)	1.45***(.345)	.709**(.293)	.947***(.140)	.944***(.139)
Landholding('00000)	.000(.000)	.002**(.001)	.000(.000)	.000(.000)	.000(.000)
Rural	-.708***(.272)	.312(.865)	.701(.582)	-.488**(.234)	-.488***(.233)
Male	-.100(.179)	-.234(.339)	.121(.302)	-.097(.140)	-.092(.140)
Married	.482**(.214)	.032(.382)	.515(.361)	.440***(.165)	.444***(.165)
Hhsize	.016(.037)	.042(.076)	.075(.059)	.033(.029)	.032(.029)
Lhdi	4.04(.376)	-6.72**(2.76)	.353(4.01)	-3.62**(1.81)	-3.28*(1.80)
Chronic illness	2.16***(.381)	1.02*(.667)	-.077(.785)	1.53***(.303)	1.82***(.320)
Not ill	-30.32(.)	-41.44(.)	-26.96(.)	-30.21(.)	-29.64(.)
Age:34-44	-.562*(.312)	.401(.602)	-.610(.494)	-.351(.242)	-.356*(.242)
Age:45-54	-.534*(.319)	.885(.618)	-.387(.504)	-.232(.248)	-.241(.248)
Age:55-64	-.118(.339)	.353(.613)	.016(.517)	.169(.257)	.118(.257)
Age:65-74	-.282(.378)	.977(.702)	-1.37*(.770)	.050(.301)	-.018(.301)
Age:≥75	-.399(.473)	.212(.792)	.972(1.06)	.042(.372)	-.018(.371)
Primary	-.254(.254)	-1.09**(.446)	.374(.401)	-.292*(.193)	-.303*(.193)
High school	-.252(.263)	-.422(.559)	-.711(.567)	-.232(.218)	-.239(.217)
University	.058(.290)	-.540(.905)	.256(1.11)	.111(.258)	.109(.257)
Illness started≤ five years	.023(.890)	-2.02(1.83)	-1.30(1.73)	-.587(.726)	-.578(.724)
Illness started≤ ten years	.354(.901)	-1.78(1.81)	-1.67(1.76)	-.474(.737)	-.473(.735)
Illness started≥ ten years	-.047(.917)	-1.86(1.88)	-2.38(1.79)	-.746(.751)	-.797(.749)
Work missed≤ month	1.78***(.361)	2.04***(.705)	2.07***(.589)	1.92***(.281)	2.01***(.281)
Work missed> month	2.49***(.338)	3.74***(.719)	3.06***(.665)	2.89***(.276)	2.91***(.275)
Matwali*Chronic illness	-	-	-	-	-.476(.366)
Matwali*Not ill	-	-	-	-	-.362(.)
Pani Nachalne*Chronic illness	-	-	-	-	-1.44***(.377)
Pani Nachalne*Not ill	-	-	-	-	-.943(.)
Matwali	-	-	-	-.734***(.189)	-.551***(.227)
Pani Nachalne	-	-	-	-.562***(.215)	-.037(.254)
Constant	-5.83***(2.17)	-9.36**(3.99)	-2.51(4.00)	-4.60***(1.67)	-4.66***(1.67)
R^2	.4544	.4595	.4791	.4536	.4541
Obs.	11055	4735	3700	19490	19490

Notes: Standard errors in parentheses. * significant at 10%, ** significant at 5% and *** significant at 1%.

-Age less than 35 years, illiterate, far-western region, illness started less than a year, work missed less than a week and *Tagadhari* are reference group respectively for age, education, region, illness started year, work missed and caste variables.

-Standard errors for not ill variable is not reported since this variable has all zeros in depended variable.

Table 4.4: Tobit regression: 2010

Dependent variable: OOP expenses					
Variable	*Tagadhari*	*Matwali*	*Pani Nachalne*	All	Interaction
	1	2	3	4	5
Lincome	.401***(.112)	.245(.193)	.332*(.171)	.356***(.085)	.356***(.085)
Landholding ('00000)	.000(.000)	.000(.000)	.000(.000)	.000(.000)	.000(.000)
Rural	-.560***(.197)	-1.09***(.478)	-.756***(.292)	-.669***(.156)	-.667***(.156)
Male	.023(.121)	-.609***(.219)	-.054(.179)	-.081(.094)	-.106(.093)
Married	.610***(.182)	.820***(.307)	.871***(.273)	.742***(.135)	.732***(.137)
Hhsize	.083***(.028)	-.013(.049)	.089**(.041)	.064***(.021)	.064***(.021)
Lhdi	-.252(.272)	.818(.761)	1.67(.573)	-.331**(.155)	.092**(.236)
Chronic illness	3.41***(.290)	5.32***(.522)	2.44***(.436)	3.63***(.224)	3.74***(.238)
Not ill	-33.84(.)	-30.05(.)	-27.76(.)	-34.80(.)	-34.77(.)
Age:34-44	-.072**(.196)	-.263(.360)	.247(.310)	-.087(.154)	-.082(.152)
Age:45-54	.512**(.204)	-.269(.386)	.298(.315)	.261*(.158)	.266*(.158)
Age:55-64	.620***(.218)	-.464*(.421)	-.527*(.345)	.115(.172)	.116(.172)
Age:65-74	.772***(.254)	-.175(.499)	-.958**(.422)	.302*(.201)	.298(.201)
Age:≥75	.575*(.314)	.468(.625)	.324(.618)	.509*(.258)	.499(.257)
Primary	-.092(.184)	.506*(.311)	-.144(.263)	.009(.138)	-.069(.139)
High school	.055(.184)	.919***(.349)	.050(.310)	.239(.146)	.234*(.146)
University	.586***(.170)	1.23***(.356)	.572*(.300)	.701***(.137)	.702***(.137)
Illness started≤ five years	-.590*(.300)	-1.78***(.522)	.002(.464)	-.688***(.234)	-.691***(.234)
Illness started≤ ten years	-.613*(.311)	-2.73***(.606)	.111(.492)	-.897***(.246)	-.904***(.246)
Illness started≥ ten years	-.886***(.321)	-2.46***(.606)	.348(.550)	-.970***(.256)	-.992***(.256)
Work missed≤ month	2.27***(.289)	3.04***(.559)	1.74***(.377)	2.32***(.222)	2.32***(.220)
Work missed> month	2.24***(.655)	4.73***(1.02)	1.68***(.614)	2.58***(.428)	2.57***(.428)
Matwali*Chronic illness	-	-	-	-	-.031(.224)
Matwali*Not ill	-	-	-	-	4.65(.)
Pani Nachalne*Chronic illness	-	-	-	-	-.511**(.239)
Pani Nachalne*Not ill	-	-	-	-	.391(.)
Matwali	-	-	-	-.829***(.121)	-.807***(.153)
Pani Nachalne	-	-	-	-.231*(.133)	-.024(.167)
Constant	-1.40(1.38)	-.751*(2.34)	1.50(2.73)	-1.09(1.06)	-1.13(1.06)
R^2	.4474	.4199	.4696	.4379	.4380
Obs.	11823	5242	3914	20979	20979

Notes: Standard errors in parentheses. * significant at 10%, ** significant at 5% and *** significant at 1%.

-Age less than 35 years, illiterate, far-western region, illness started less than a year, work missed less than a week and *Tagadhari* are reference group respectively for age, education, region, illness started year, work missed and caste variables.

-Standard errors for not ill variable is not reported since this variable has all zeros in depended variable.

Table 4.5: Tobit regression: Marginal effects

Dependent variable: OOP expenses				
Variable	*Year: 2003*		*Year: 2010*	
	1	2	3	4
Lincome	.021***(.003)	.022***(.003)	.007***(.001)	.009***(.002)
Landholding ('00000)	.000(.000)	.000(.000)	.000(.000)	.000(.000)
Rural	-.011***(.004)	-.012***(.005)	-.014***(.003)	-.016***(.004)
Male	-.002(.003)	.002(.003)	-.002(.002)	-.002(.002)
Married	.010***(.003)	.010***(.004)	.015***(.003)	.017***(.003)
Hhsize	.001(.001)	.001(.001)	.001***(.000)	.002***(.000)
Lhdi	-.066**(.029)	-.066***(.029)	.002**(.005)	.002(.005)
Chronic illness	.038***(.008)	.07***(.008)	.088***(.006)	.106***(.007)
Not ill	-3.68***(.049)	-3.55***(.048)	-4.41***(.039)	-5.27***(.044)
Age:34-44	-.008*(.005)	-.008*(.005)	-.001(.003)	-.002(.003)
Age:45-54	-.005(.005)	-.005(.005)	.005*(.003)	.006*(.004)
Age:55-64	.004(.006)	.002(.006)	.002(.003)	.002(.004)
Age:65-74	.001(.007)	.000(.007)	.006*(.004)	.007*(.005)
Age:≥75	.001(.009)	-.001(.009)	.011*(.005)	.012*(.006)
Primary	-.006*(.004)	-.007*(.004)	.000(.002)	.000(.003)
High school	-.005(.005)	-.005(.005)	.005*(.003)	.005*(.003)
University	.002(.006)	.002(.006)	.015***(.003)	.017***(.003)
Illness started≤ five years	-.013(.016)	-.014(.016)	-.014***(.004)	-.016***(.005)
Illness started≤ ten years	-.011(.017)	-.011(.017)	-.018***(.004)	-.021***(.005)
Illness started≥ ten years	-.018(.019)	-.020(.019)	-.020***(.005)	-.023***(.005)
Work missed≤ month	.050***(.008)	.053***(.008)	.055***(.005)	.063***(.006)
Work missed> month	.078***(.008)	.080***(.008)	.062***(.011)	.071***(.013)
Matwali*Chronic illness	-	-.012(.008)	-	-.001(.005)
Matwali*Not ill	-	-.009***(.000)	-	.115***(.002)
Pani Nachalne*Chronic illness	-	-.033***(.008)	-	-.012***(.005)
Pani Nachalne*Not ill	-	-.022(.000)	-	.086***(.001)
Matwali	-.018***(.005)	-.014**(.005)	-.017***(.002)	-.019***(.003)
Pani Nachalne	-.013***(.005)	-.001(.006)	-.002**(.002)	-.001(.004)

Notes: Standard errors in parentheses. * significant at 10%, ** significant at 5% and *** significant at 1%.

-Age less than 35 years, illiterate, far-western region, illness started less than a year, work missed less than a week and *Tagadhari* are reference group respectively for age, education, region, illness started year, work missed and caste variables.

Table 4.6: Ordered probit estimation: 2003

Dependent variable: SAH categorical variable				
Variable	*Very Poor*	*Poor*	*Fair*	*Good*
OOP ratio	.0000(.0000)	.000(.000)	.007(.006)	-.007(.006)
Lnholding('00000)	-.0000(.0000)	-.000(.000)	-.000(.000)	.000(.000)
Rural	.0001***(.0000)	.003***(.000)	.104***(.010)	-.107***(.010)
Male	-.0000(.0000)	-.001***(.000)	-.022***(.007)	.023***(.007)
Married	-.0000(.000)	-.001***(.000)	-.022***(.008)	.023***(.008)
Hhsize	.0000(.000)	.001***(.000)	.011***(.001)	-.011***(.001)
Lhdi	.0000(.0000)	.000(.000)	.011(.011)	-.012(.012)
Chronic illness	.0009**(.0004)	.036***(.004)	.294***(.023)	-.331***(.030)
Not ill	-.0001***(.0000)	-.006***(.000)	-.115***(.010)	.121***(.011)
Age:35-44	.0000(.0000)	.002***(.000)	.055***(.012)	-.058***(.013)
Age:45-54	.0001***(.000)	.006***(.001)	.112***(.013)	-.119***(.014)
Age:55-64	.0002***(.000)	.012***(.001)	.169***(.013)	-.182***(.015)
Age:65-74	.0003***(.0001)	.015***(.002)	.193***(.017)	-.209***(.019)
Age≥75	.0006***(.0002)	.026***(.004)	.248***(.019)	-.275***(.023)
Primary	-.0000(.0000)	-.002***(.000)	-.067***(.008)	.069***(.009)
High school	-.0000(.0000)	-.003***(.000)	-.085***(.010)	.088***(.010)
University	-.0001***(.0000)	-.004***(.000)	-.117***(.012)	.120***(.012)
Illness started≤ five years	.0000(.0000)	.003(.005)	.058(.084)	-.061(.090)
Illness started≤ ten years	.0000(.0000)	.004(.006)	.082(.085)	-.082(.092)
Illness started> ten years	.0000(.0000)	.001(.002)	.050(.085)	-.055(.087)
Work missed≤month	.0001*(.00005)	.004**(.002)	.085***(.032)	-.090***(.034)
Work missed>month	.0007**(.0002)	.028***(.006)	.254***(.024)	-.283***(.030)
Matwali	.0001(.0001)	.001(.001)	.016*(.008)	-.017*(.009)
Pani Nachalne	.0001***(.0000)	.004***(.001)	.080***(.009)	-.084***(.010)
Log likelihood		-12857.26		
Pseudo R^2		.1257		
Obs.		19490		
Likelihood ratio test:		$\chi^2(53) = 719.71$ P value=000		

Notes: Standard errors in parenthesis. * significant at 10%, ** significant at 5% and *** significant at 1%.

-Other illness, age less than 35 years, illiterate, far-western region, illness started less than a year, work missed less than a week and *Tagadhari* are reference group respectively for illness, age, education, region, illness started year, work missed and caste variables.

-Likelihood ratio test for identical coefficients for explanatory variables across subsamples.

Table 4.7: Ordered probit estimation: 2010

Dependent variable: SAH categorical variable				
Variable	*Poor*	*Fair*	*Good*	*Excellent*
OOP ratio	.0000(.0000)	.001***(.000)	.011***(.003)	-.011***(.003)
Lnholding('00000)	-.0000(.0000)	-.000(.000)	-.000(.000)	.000(.000)
Rural	-.0000(.0000)	-.001(.001)	-.018*(.009)	.019*(.009)
Male	-.0000(.0000)	-.000(.000)	-.005(.006)	.005(.007)
Married	-.0000(.0000)	-.000(.000)	-.008(.009)	.009(.009)
Hhsize	.0000(.0000)	.000(.000)	.001(.001)	-.001(.001)
Lhdi	-.0001***(.000)	-.004***(.000)	-.138***(.010)	.142***(.010)
Chronic illness	.0047***(.0013)	.086***(.011)	.394***(.009)	-.485***(.019)
Not ill	-.0001***(.0000)	-.002***(.000)	-.048***(.009)	.050***(.010)
Age:35-44	.0001**(.00002)	.001***(.000)	.042***(.011)	-.044***(.011)
Age:45-54	.0001***(.0000)	.004***(.000)	.083***(.012)	-.088***(.013)
Age:55-64	.0001***(.0000)	.007***(.001)	.120***(.013)	-.128***(.015)
Age:65-74	.0003***(.0001)	.012***(.002)	.169***(.016)	-.182***(.017)
Age≥75	.0008***(.0002)	.025***(.004)	.244***(.017)	-.271***(.021)
Primary	-.0001***(.0000)	-.002***(.000)	-.074***(.009)	.077***(.010)
High school	-.0001***(.0000)	-.002***(.000)	-.069***(.010)	.072***(.010)
University	-.0001***(.0000)	-.003***(.000)	-.074***(.010)	.077***(.010)
Illness started≤ five years	-.0000(.0000)	-.002**(.001)	-.048*(.031)	.050*(.032)
Illness started≤ ten years	-.0000(.0000)	-.001(.001)	-.041(.033)	.043(.034)
Illness started> ten years	-.0000(.0000)	-.001(.001)	-.011(.034)	.011(.036)
Work missed≤month	.0001(.0001)	.003**(.002)	.072**(.029)	-.076**(.031)
Work missed>month	.0013*(.0007)	.036***(.013)	.275***(.034)	-.313***(.047)
Matwali	-.00004***(.00001)	-.002***(.000)	-.064***(.008)	.066***(.008)
Pani Nachalne	.00002**(.00001)	.001**(.0004)	.019**(.009)	-.020**(.009)
Log likelihood		-12459.4		
Pseudo R^2		.2323		
Obs.		20979		
Likelihood ratio test:		$\chi^2(53) = 384.54$ P value=000		

Notes: Standard errors in parenthesis. * significant at 10%, ** significant at 5% and *** significant at 1%.

-Age less than 35 years, illiterate, far-western region, illness started less than a year, work missed less than a week and *Tagadhari* are reference group respectively for age, education, region, illness started year, work missed and caste variables.

-Likelihood ratio test for identical coefficients for explanatory variables across subsamples.

Table 4.8: Generalized ordered probit estimation: 2003

Dependent variable: SAH categorical variable			
Variable	*Model 1*	*Model 2*	*Model 3*
OOP ratio	.000(.000)	.000(.000)	.080**(.039)
Lnholding('00000)	-.000(.000)	.000(.000)	-.000(.000)
Rural	-.000(.000)	.001(.001)	.085***(.013)
Male	.000(.000)	.000(.000)	-.031***(.007)
Married	-.000(.000)	-.001(.001)	-.027***(.009)
Hhsize	.000(.000)	.000(.000)	.011***(.001)
Lhdi	-	.955***(.049)	-.214(.056)
Chronic illness	.001(.001)	.005**(.003)	.430***(.038)
Not ill	-.000(.000)	-.025***(.003)	-.088***(.012)
Age:35-44	.000(.000)	.001(.002)	.076***(.014)
Age:45-54	-.000(.000)	.002(.001)	.139***(.015)
Age:55-64	-.000(.000)	.006***(.002)	.218***(.017)
Age:65-74	-.000(.000)	.005***(.002)	.271***(.021)
Age≥75	-	.018***(.006)	.323***(.027)
Primary	-.000(.000)	-.001***(.000)	-.059***(.009)
High school	-.000(.000)	-.002***(.000)	-.077***(.011)
University	-	-	-.127***(.013)
Illness started≤ five years	.889***(.202)	-.001(.002)	.062(.112)
Illness started≤ ten years	.895***(.204)	-.001(.002)	.133(.116)
Illness started> ten years	.000(.000)	-.002(.005)	.070(.119)
Work missed≤month	-	.003*(.002)	.150***(.058)
Work missed>month	.007***(.011)	.028***(.008)	.199***(.063)
Matwali	-.000(.000)	-.001(.001)	.012(.010)
Pani Nachalne	-.000(.000)	.001(.001)	.045***(.012)
Log likelihood	-61.14	-1164.26	-10722.59
Pseudo R^2	.5273	0.3387	.1861
Obs.	4747	18585	19490

Notes: Standard errors in parentheses. * significant at 10%, ** significant at 5% and *** significant at 1%.

-Age less than 35 years, illiterate, far-western region, illness started less than a year, work missed less than a week and *Tagadhari* are reference group respectively for age, education, region, illness started year, work missed and caste variables.

-*Model 1* estimates a binary probit model where dependent variable being 1 for *Very poor* SAH category and zero otherwise.

-*Model 2* estimates a binary probit model where dependent variable being 1 for *Very poor* and *Poor* SAH categories and zero otherwise.

-*Model 3* estimates a binary probit model where dependent variable being 1 for *Very poor, Poor* and *Fair* SAH categories and zero otherwise.

Table 4.9: Generalized ordered probit estimation: 2010

Dependent variable: SAH categorical variable			
Variable	*Model 1*	*Model 2*	*Model 3*
OOP ratio	.000(.000)	.001***(.000)	.013(.009)
Lnholding('00000)	.0000(.0000)	-.000(.000)	-.000(.000)
Rural	.000(.000)	-.002**(.001)	-.045***(.012)
Male	.000(.000)	.001(.001)	-.010(.007)
Married	-.002**(.001)	-.005***(.001)	-.003(.010)
Hhsize	-.000(.000)	-.000(.000)	.002**(.001)
Lhdi	-.001(.001)	-.001(.001)	-.110***(.020)
Chronic illness	.011(.012)	.046***(.011)	.522***(.025)
Not ill	.000(.000)	-.012***(.002)	-.065***(.010)
Age:35-44	-.001(.001)	.003***(.001)	.051***(.013)
Age:45-54	-.000(.000)	.001(.001)	.123***(.015)
Age:55-64	.000(.000)	.007***(.002)	.147***(.017)
Age:65-74	.004**(.002)	.012***(.004)	.202***(.021)
Age≥75	.002(.003)	.034***(.008)	.288***(.026)
Primary	-.001**(.0005)	-.003***(.001)	-.063***(.011)
High school	-.001**(.0005)	-.004***(.001)	-.051***(.012)
University	-.001**(.0005)	-.005***(.001)	-.063***(.011)
Illness started≤ five years	-.001(.001)	-.004***(.001)	-.010(.049)
Illness started≤ ten years	-.001**(.0004)	-.003***(.001)	-.008(.052)
Illness started> ten years	-.000(.000)	-.002**(.001)	.038(.056)
Work missed≤month	.005(.008)	.010***(.004)	.023(.037)
Work missed>month	.004(.009)	.056***(.023)	.198**(.079)
Matwali	.000(.000)	.001(.001)	-.019*(.010)
Pani Nachalne	-.001**(.0002)	.001(.001)	.032**(.011)
Log likelihood	-118.38	-1519.22	-10646.01
Pseudo R^2	.2032	.3051	.2577
Obs.	5419	20662	20979

Notes: Standard errors in parentheses. * significant at 10%, ** significant at 5% and *** significant at 1%.

-Age less than 35 years, illiterate, far-western region, illness started less than a year, work missed less than a week and *Tagadhari* are reference group respectively for age, education, region, illness started year, work missed and caste variables.

-*Model 1* estimates a binary probit model where dependent variable being 1 for *Poor* SAH category and zero otherwise.

-*Model 2* estimates a binary probit model where dependent variable being 1 for *Poor* and *Fair* SAH categories and zero otherwise.

-*Model 3* estimates a binary probit model where dependent variable being 1 for *Poor, Fair* and *Good* SAH categories and zero otherwise.

Chapter 5

Discussion

This chapter synopsises the methodological and empirical findings generated by this dissertation. As this dissertation is composed of three different empirical studies, findings together with the limitation are discussed separately and agendas for future researches are explored.

5.1 Chapter 1

This study investigates the parental plan regarding daughters' age at marriage and its effects on their decision to invest in daughter's education. Unlike other studies that argue parental anti-bias in female education and child marriage as contributing factors to undermine female education this study concentrates on indirect disincentive to female education that marriage exerts via implied division of household labour and this applies even to the female who marries post-childhood.

This study contributes to the literature by developing a theoretical framework for jointly determining female education and planned age at marriage. The framework is based on Jafarey (2011), in which gender wage inequality is shown to lead to both a direct discount on female education and an indirect one following from the

marital division of labour which allocates women to spend relatively more time in housework and men in market work. It is shown that the earlier the planned age at marriage the lesser incentive parents will have in investing in daughter's education. In addition, this framework acknowledges that the age at marriage can itself depend on individual and cultural factors, such as a female's ability to benefit from schooling and/or cultural expectations regarding an ideal age for her to marry.

This hypothesis is tested using household survey data from Nepal. Empirical evidence validates the hypothesis that female education is negatively affected by cultural practices that favours early marriage: increasing female marriage by one year would produce, on average, 0.4 year increment in schooling.

Addressing endogeneity

As theoretical framework developed in this study shows that a female may select into early marriage on the basis of idiosyncratic and unobservable differences in ability, an ordinary least square method produces bias results and thus an instrumental variable estimation strategy is applied to estimate the casual effect of age at marriage on female education. Therefore, potential endogeneity is controlled for by exploiting variations in cultural norms regarding dowry and differences in average age of female's marriage among ethnicities and regions as instrumental variables.

The validity of both IVs is discussed referring to available theoretical arguments as well as empirical evidence. A statistical test has also been carried out for the validity of IVs.

Limitation of this study

Theoretical framework outlined in this study suggests that girl's parents invest on her education depending on her ability to benefit from education. That is, if parents

judge her to be of relatively low ability they may decide both not to school her to a very high level far, and to make better use of her time or to marry her at an early age compared to other girls within her community. This proposition is not tested explicitly due to the data limitation. It is not possible to detect parent's household from the survey data used in this analysis. Future research should take this issue into account. However, it has been argued in theoretical section that parents invest in daughter's education depending on her ability to acquire and benefit from education and they marry her early if parents judge her to be of low ability. And thus an instrumental variable strategy is used to estimate casual effect of age at marriage on female education. Empirical results show larger values of age at marriage coefficients compared to ordinary least square estimates in all cases. It implies that ability is positively correlated to age at marriage.

5.2 Chapter 2

This study evaluates the sources of caste wage differentials arguing that caste-based division of labour can perpetuate itself through the inter-generational transmission of low level of educational endowment in low castes and caste differences in access to large firms and better occupations. Oaxaca (1973) decomposition is widely used methodology in disentangling pre-market and current market effects on wage differentials which accounts for productivity characteristics as endowments while estimating such effects. This methodology was later expanded by Banerjee and Knight (1985) to include occupation as an additional endowment.

This study goes further in capturing the effect on wage inequality by introducing firm characteristics to supplement educational and occupational differences. Therefore, an interaction model of decomposition is introduced to see the effect of caste differences in access to larger firms on caste wage differentials. Two waves of na-

tionally representative household survey data from Nepal are used.

Empirical results show that caste wage inequality is present in Nepalese labour market in both periods which is increasing over time particularly for the lowest caste. Furthermore, results show that caste differences in human capital endowments are important for explaining wage inequality, but also so are occupational and firm size effects, especially when the latter two are taken together.

The main methodological contribution of this study is to use an expanded set of proxies for decomposing gross wage differentials into the pre-market and current market effects.

Limitation of this study

This study excludes a considerable portion of wage workers while estimating the source of caste wage differentials in the labour market. Reason behind is that the majority of wage earners in Nepal are from agricultural labourers who did not report the size of their employers. Additionally, those working as subsistence agricultural labourers may not be relevant in estimating the effect of access to larger firms on wage differentials as this study intends to. Therefore, future work that estimates expanded decomposition methodology proposed in this study, which includes firm characteristics along with occupational distribution and productivity characteristics as endowments, employing data from different labour market will be an additional advantage to generalise the finding of this study and to make it more robust.

5.3 Chapter 3

This study argues that historically discriminated-against castes inherit low levels of social capital in low castes. These variables can have negative effects on the

utilisation of health care services which in turn may transmit into the inferior health outcomes in low castes.

Findings from this study suggest that caste-inequity in health care utilisation as well as in health outcomes exist in Nepal. However, in indicates that both types of inequity are decreasing over time. More importantly, empirical evidence suggests that caste differences in health care utilisation is coming from the caste differences in household income along with need variables. It indicates that policies that promote income equality can positively contribute to the equity in health care utilisation across castes.

Limitation

Although this study incorporates caste context into the analysis of health sector inequity which is distinct from other varieties of social classification such as race, gender and ethnicity, it has its own limitation. Firstly, health care utilisation in this study is proxied by OOP payment which may not accurately represent the utilisation of health care services both in quantitative and qualitative terms. Secondly, two waves of survey data used in this study did not include the same respondents which may caution to conclude about the trend of caste inequity in health care utilisation over time. Future researches should take these issues into account. However, this study estimates the impact of caste on health outcomes which has not yet been carried out, at least to my knowledge and contributes to the literature of health economics.

Bibliography

Agee, M.D. and T. D. Crocker (1996). Parents' Discount Rates for Child Quality. *Southern Economic Journal 63 (1), 36-50*

Akerlof, G. A. (1984). *An Economic Theorist's Book of Tales.* Cambridge: Cambridge University Press.

Alder, N. E. and K. Newman (2002). Socioeconomic Disparities in Health: Pathways and Policies. *Health Affairs 21(2), 60-70.*

Altonji, G. J. and R. M. Blank (1999). Race and gender in the labor market. In *Handbook of Labor Economics.* Amsterdam: Elsevier.

Anderson, S. (2004). Dowry and Property Rights. Website. `http://ipl.econ.duke.edu/bread/papers/working/080.pdf`

Anne, C. and A. Deaton (2005). Health and Wealth among the poor: India and South Africa Compared. *Americal Economic Review 95(2), 229-233.*

Arrow, J. K. (1971). The Theory of Discrimination. Website. `http://www2.econ.iastate.edu/classes/econ321/rosburg/Arrow%20-%20The%20Theory%20of%20Discrimination.pdf`

Asadullah, M. N. ((2006). Returns to education in Bangladesh. Wbsite. `http://www3.qeh.ox.ac.uk/pdf/qehwp/qehwps130.pdf`

Aslam, M. (2009). Education Gender Gaps in Pakistan: Is the Labor Market to Blame?. *Economic Development and Cultural Change 57(4), 747-784.*

Atal, J. P., H. Nopo and N. Winder (2009). New Centuries Old Disparities: Gen-

der and Ethnic Wage Gaps in Latin America. Website. http://idbdocs.iadb.org/wsdocs/getdocument.aspx?docnum=2230955

Autor, D. H.(2003). Lecture Notes: Economics of Discrimination-Theory. MIT 14.661. Website. http://econ-www.mit.edu/file/554.

Babu, B.V., G.P. Chhotray, R.K. Hazra, K. Satyanarayana (2001). Community Perception of a District Health System. *Journal of Health Management 3 (1), 1-13.*

Banerjee, A., A. Deaton and E. Duflo (2004). Wealth, Health and Health Service in Rural Rajasthan. *American Economic Review 94(2), 326-330.*

Banerjee, B. and J. B. Knight (1985). Caste discrimination in the Indian urban labour market. *Journal of Development Economics 17(3), 277-307.*

Baru, R., A. Acharya, S. Acharya, A K., Shivakumar and K. Nagraj (2010). Inequities in Access to Health Services in India: Caste, Class and Region. *Economic and Political Weekly 38, 49-58.*

Bayisenge, J. (2010). Early marriage as a barrier to girl's education: A development challenge in Africa. Website. http://www.ifuw.org/fuwa/docs/Early-marriage.pdf

Becker, G. and N. Tomes (1976). Child Endowments and the Quantity and Quality of Children. *The Journal of Political Economy 84 (4), 143-162.*

Becker, G. S. (1957). *The Economics of Discrimination.* Chicago: University of Chicago Press.

Becker, G. S. (1962). Investment in Human Capital: A Theoretical Analysis. *The Journal of Political Economy 70(5),9-49.*

Becker, G. S. (1971). *The Economics of Discrimination* 2ed edition. Chicago: University of Chicago Press.

Becker, G. S. (1973). A Theory of Marriage: Part I. *The Journal of Political Economy 81(4), 813-846.*

Becker, G. S. (1974). A Theory of Marriage: Part II. *The Journal of Political Econ-*

omy 82(2), S11-S26.

Becker, G. S. (1974). A Theory of Social Interactions. *The Journal of Political Economy 82 (6),1063-1093.*

Becker, G. S. (1975). *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education, 2ed Edition.* Chicago: University Press of Chicago press.

Behrman, J. R. and A.B. Deolalikar (1995). Are There Differential Returns to Schooling by Gender? The Case of Indonesian Labour Markets. *Oxford Bulletin of Economics and Statistics 57(1), 97-117.*

Behrman, J. R., R. A. Pollak and P. Toubman (1982). The Parental Preferences and Provision for Progeny. *Journal of Political Economy 90(1), 52-73.*

Ben-Porath, Y. (1967). The Production of Human Capital and the Life Cycle of Earnings. *The Journal of Political Economy 75(4), 352-365.*

Bennett, L. (2005). Gender, Caste and Ethnic Exclusion in Nepal: Following the Policy Process from Analysis to Action. Conference paper.

Website. `http://www.k4health.org/sites/default/files/Gender`

Bennett, L., D. R. Dhital and P. Govindasamy (2008). Caste, Ethnic and Regional Identity in Nepal: Further Analysis of the 2006 Nepal Demographic and Health Survey. Website. `http:www.measuredhs.com/pubs/pdf/FA58/FA58.pdf`

Berdahl, T. A., J. B. Kirby and R. A. Torres Stone (2007). Access to Health Care for Nonmetro and Metro Latinos of Mexican Origin in the United States. *Medical Care 45(7), 647-654.*

Billari, F. C., A. Prskawetz, and J. Furnkranz (2002). The cultural evolution of of age-at-marriage norms. Max Planck Institute for Demographic Re- search Working Paper WP-2002-018. `http://www.demogr.mpg.de/papers/working/wp-2002-018.pdf.`

Birdsall, N. and J. L. Londono (1997). Assets Inequality Matters: An Assessment of

the World Bank's Approach to Poverty Reduction. *The American Economic Review 87(2), 32-37.*

Blinder, A. S. (1973). Wage Discrimination: Reduced Form and Structural Estimates. *The Journal of Human Resources 8(4), 436-455.*

Bond, G. C. (1981). Social Economic Status and Educational Achievement: A Review Article. *Anthropology Education Quarterly 12(4), 227-257.*

Boskin, M. j.(1974). A Conditional Logit Model of Occupational Choice. *Journal of Political Economy 82, 389-398.*

Botwinic, H. (1993). *Persistent Inequalities: Wage Disparity under Capitalist Competition.* New York: Princeton University Press.

Braveman, P. and S. Gruskin (2003). Poverty, equity, human rights and health. *Bulletin of the World Health Organization 81 (7), 539-545.*

Brein, J. Michael and Lee A. Lillard (1994). Education, Marriage, and First Conception in Malaysia. *The Journal of Human Resources 29(4), 1167-1204.*

Brown, C. J. Medoff (1989). The Employer Size-Wage Effect. *The Journal of Political Economy 97(5), 1027-1059.*

Brown, R. S., M. Moon and B. S. Zoloth (1980). Incorporating Occupational Attainment in Studies of Male-Female Earning Differentials. The Journal of Human Resources 15(1), 3-28.

Cain, G. G. (1984). The economics of discrimination: Part 1. *Focus 7(2), 1-11.*

Cameron, L. A. and C. Worswick (2001). Education Expenditure Responses to Crop Loss in Indonesia: A Gender Bias. *Economic Development and Cultural Change 49(2), 351-363.*

Carneiro, P., J.J. Heckman and D.V. Masterov (2005). Labor Market Discrimination and Racial Differences in Premarket Factors. *Journal of Law and Economics 48(1) 1-39.*

Chase, R.S. (2000). Labor Market Discrimination During Post-communist Transition: A Monopsony Approach to the Status of Latvia's Russian Minorities. Website. http://deepblue.lib.umich.edu/bitstream/handle/2027.42/39765/wp381.pdf?sequence

Chaturvedi, N. (2003). Ethnic Differences in Cardiovascular Disease. *Heart 89(6), 681-686.*

Checchi, D. (1997). Education and Intergenerational Mobility in Occupations: A Comparative Study. *The American Journal of Economics and Sociology 56(3), 331-351.*

Chen, A. Y. and J. J. Escarce (2004). Quantifying income-related inequality in healthcare delivery in the United States. *Medical Care 42(1), 38-47.*

Chiswick, B. (1970). "An Interregional Analysis of Schooling and the Skewness of Income" in Hansen Lee. W. (ed) *Education, Income and Human Capital.* New York. National Bureau of Economic Research.

Cochrane, S. H. (2009). Fertility and Education: What do We Really Know?. Website. http://documents.worldbank.org/curated/en/1979/01/1558317/fertility-education-rea

Cook, B L., T.G. McGuire, K. Lock and A. M. Zaslavsky (2010). Comparing methods of racial and ethnic disparities measurement across different settings of mental health care. *Health Service Research-HSR 45(3), 825-847.*

Cox, O. C. (1944). Class and Caste: A Definition and a Distinction. *The Journal of Negro Education 13(2), 139-149.*

Cox, T. (1988). A Current Socio-economic Status of Untouchable in Nepal. *Occasional Papers on Sociology and Anthropology 4, 90-109.*

Dahal, D. R., T. Fricke and A. Thornton (1993). The Family Contexts of Marriage Timing in Nepal. *Ethnology 32(4), 305-323.*

Dalmia, S. and P. Lawrence (2005). The Institution of Dowry in India: Why It Continues to Prevail. *The Journal of Developing Areas 38(2), 71-93.*

D'Anna, L. H. , N. A. Ponce and J. M. Siegel (2010). Racial and ethnic health disparities: evidence of discrimination's effects across the SEP spectrum. *Ethnicity and Health 15(2), 121-143.*

Darity, W. A. and P. L. Mason (1998). Evidence on Discrimination in Employment: Codes of Colour and Codes of Gender. *The Journal of Economic Perspectives 12(2), 63-90.*

Darity, W. A. and R. M. Williams (1985). Peddlers Forever?: Culture, Competition, and Discrimination. *The American Economic Review 75(2), 256-261.*

Das, M. B. and P. V. Dutta (2007). Does Caste Matter for the Wages in the Indian Labor Market?. Website. `http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.190.3708&rep=rep1&type=pdf`

Das, S. K. (2009). Exploration of Causes of Dowry System in Dhanusha District of Nepal. Website. `http://www.socialinclusion.org.np/researchdetail-238.html`

Dasgupta, I. and D. Mukherjee (2003). 'Arranged' Marriage, Dowry and Female Literacy in a Transitional Society. Wbsite. `http://papers.ssrn.com/sol3/papers.cfm?abstract_id=463580`

Deaton, A. (2003). Health, Income and Inequality. Website. `http://www.nber.org/reporter/spring03/health.html`

Deaton, A. (2006). Global Patterns of Income and Health: Facts, Interpretation, and Policies. NBER Working Paper No. 12735. Website. `http://www.nber.org/papers/w12735`

Deaton, A. (2011). What does the empirical evidence tell us about the injustice of inequalities? Princeton University, Working Paper No. 1284. Website. `https://www.princeton.edu/rpds/papers/What_does_the_empirical_evidence_tell_us_about_the_injustice.pdf`

Deaton, A. and D. Lubotsky (2003). Mortality, Inequality and Race in American

cities and states. *Social Science and Medicine 56(6), 1139-1153.*

Deshpande, A. (2011). *The Grammar of Caste: Economic Discrimination in Contemporary India.* New Delhi: Oxford University Press.

Dhital, R. (2012). Child Marriage in Nepal. Website. `http://www.volunteeraidnepal.org/child-marriage-nepal`

Dommaraju, P. V. Agadjanian and S. Yabiku (2008). The Pervasive and Persistence Influence of Caste on Child Mortality in India. *Population Research and Policy Review 27(4), 477-495.*

Doss, C. R. (1996). Testing among models of intrahousehold resource allocation. *World Development 24(10), 1597-1609.*

Duman, A. (2010). Female education inequality in Turkey: factors affecting girls' schooling decisions. *International Journal of Education Economics and Development 1(3), 243-258.*

Ermisch, J. and M. Francesconi (2001). Family Matters: Impacts of Family Background on Educational Attainments. *Econometrica 68(270), 137-156.*

Field, E. A. Ambrus (2008). Early Marriage, Age of Menarche and Female Schooling Attainment in Bangladesh. *Journal of Political economy 116(5), 881-930.*

Garimoi, C. (2009). Health equity: challenges in low income countries. *African Health Sciences 9(S2), 49-51.*

George, S. and M. Akiko (1999). Health care financing and delivery in developing countries. *Health Affairs 18(3), 193-205.*

Ghosh, S. (2011). Catastrophic Payments and Impoverishment Due to Out-of-Pocket Health Spending. *Economic and Political Weekly 47, 63-70.*

Glewwe, P. and H. Jacoby (1994). Student Achievement and Schooling Choice in Low-Income Countries: Evidence from Ghana. *The Journal of Human Resources 29(3), 841-864.*

Greene, W., M.N. Harris, B. Hollingsworth and P. Maitra (2008). A Bivariate Latent Class Correlated Generalized Ordered Probit Model with an Application to Modeling Observed Obesity Levels. Website. http://ssrn.com/abstract=1281910

Greenwood, J., A. Seshadri and M. Yorukoglu (2005). Engines of Liberation. *Review of Economic Studies 72(1), 109-133.*

Grimshaw, D. and J. Rubery (2002). The adjusted gender pay gap: A critical appraisal of standard decomposition Techniques. Website. https://research.mbs.ac.uk/european-employment/Portals/0/docs/gendersocial/GPGpoliticalpdf

Grytten, J., G. Rongen and R. Sorensen (1995). Can a Public Health Care System Achieve Equity? The Norwegian Experience. *Medical Care 33(9), 938-951.*

Gurung, H. (2002). Trident and Thunderbolt, Cultural Dynamics in Nepalese Politics. In *Ethnicity, Caste and a Pluralist Society, ed. B. Krishna Battachan.* Patan, Nepal: Social Science Baha Himal Association.

Halli, S. (2003). Marriage Patterns in Rural India: Influence of Sociocultural Context. Website. http://paa2003.princeton.edu/abstracts/62109

Harris, K. M. , M. J. Edlund and S. Larson (2005). Racial and ethnic differences in the mental health problems and use of mental health care. *Medical Care 43(8), 775-784.*

Hettler, P. (2007). The Decomposition of Firm-size Wage Differentials. *Journal of Labor Research 28(3), 477-486.*

Hinks, T. and D. Watson (2001). A multinomial logit nondiscriminatory approach to estimating racial wage and occupational discrimination. *Applied Economics 33(5), 605-612.*

Hofer, A. (1979). *The Caste Hierarchy and the State in Nepal: A Study of the Muluki Ain of 1854.* Patan, Nepal: Himal Books.

Honda, C. and Y. Ohkusa (2003). Horizontal Inequity in Health Care Utilization in

Japan. *Health care Management Science 6 (3), 189-196.*

Hotchkiss, D. R., J. R. Jeffrey, K. Karmacharya and P. Sangraula (1998). Household health expenditures in Nepal: implications for health care financing reform. *Health Policy and Planning 13(4), 371-383.*

Hummer, R.A. (1996). Black-White Differences in Health and Mortality: A Review and Conceptual Model. *The Sociological Quarterly 37(1), 105-125.*

Inequality Trust (2013). Inequality and Health. Website. `http://inequality.org/inequality-health.`

International Monetary Fund (2003). Nepal : Poverty Reduction Strategy Paper, IMF Country Report No. 03/305. Website. `http://www.imf.org/external/pubs/cat/longres.aspx?sk=16894.0`

Jacoby, H. G. and E. Skoufias (1997). Risk, Financial Markets, and Human Capital in a Developing Country. *The Review of Economic Studies 64(3), 311-335.*

Jafarey, S. (2011). Glass Slippers and Glass Ceiling: An analysis of Marital Anticipation and Female Education. Website. `https://manageweb.ict.uniba.it/ricerca/dipartimenti/dse/seminari/seminari-2011/jefarey2011.pdf`

Jann, B. (2008). The Blinder-Oaxaca decomposition for linear regression models . *The Stata Journal 8(4), 453-479.*

Jeffrey, C. (2001). 'A Fist Is Stronger than Five Fingers': Caste and Dominance in Rural North India. *Transactions of the Institute of British Geographers New Series 26(2), 217-236.*

Kawachi, I. , N. Daniels and D. E. Robinson (2005). Health Disparities by Race and Class. Why Both Matter? *Health Affairs 24 (2), 343-353.*

Keiko, A. and T. Yoshinori (2006). Socio Cultural Factors Affecting Girls' Limited Access to School Education in North West Frontier Province in Pakistan. Website. `http://edisdat.ied.edu.hk/pubarch/b15907314/full_paper/1394178079.pdf`

King, E.M. and A.D. Mason (2001). Engineering Development Through Gender Equality in Rights, Resources and Voice. Website. http://siteresources.worldbank.org/PGLP/Resources/Engendering_Development.pdf

King, E.M. and M. A. Hill (1993). *Women's Education in Developing Countries: Barriers, Benefits and Policies.* Baltimore: The John Hopkins University Press.

Kingdom, G. (1998). Does the labour market explain lower female schooling in India? *The Journal of Development Studies 35(1), 39-65.*

Kirby, J. B. and T. Kaneda (2005). Neighborhood Socioeconomic Disadvantage and Access to Health Care. *Journal of Health and Social Behavior 46(1), 15-31.*

Krieger, N. (1987). Shades of Difference: Theoretical Underpinnings of the Medical Controversy on Black/White Differences in the United States, 1830 - 1870. *International Journal of Health Services 17(2), 259-278.*

Kuran, T. (1987). Preference Falsification, Policy Continuity and Collective Conservatism. *The Economic Journal 97 (387), 642-665.*

Lahiri, S. and S. Self (2007). Gender Bias in Education: the Role of Inter-household Externality, Dowry and other Social Institutions. *Review of Development Economics 11 (4), 591-606.*

Leon, D. and G. Walt (2001). *Poverty, Inequality, and Health: An International Perspective.* New York: Oxford University Press.

Lerner, M., B. Candida and R. Hisrich (1997). Israeli Women Entrepreneurs: An Examination of Factors Affecting Performance. *Journal of Business Venturing 12(4), 315-339.*

Lerrea, C. and F. M. Torres (2006). *Indigenous Peoples, Poverty and Human Development in Latin America Edited by G. Hall and H. A. Patrinos.* London: Palgrave Macmillan.

Liu, Y., W.C. Hsiao and K. Eggleston (1999). Equity in health and health care: the

Chinese experience. *Social Science and Medicine 49(10), 1349-1356.*

Lordan, G., E. G. Soto, R.P. C., Brown and I. Correa-Valez (2012). Socioeconomic status and health outcomes in a developing country. *Health Economics 21(2), 178-186.*

Lovell, P. A. (2006). Race, Gender and Work in Sao Paulo, Brazil, 1960-2000. *Latin American Research Review 41(3), 63-87.*

Lu. J. R., G. M., Luenge, S. Kwon, Keith Y.K., Tin, E. van Doorslaer, O. O'donell (2007). Horizontal equity in health care utilization evidence from three high-income Asian economies. *Social Science and Medicine 64, 199-212.*

Madden, J. F. (1975). Discrimination- A Manifestation of Male Market Power?. In *Sex Discrimination and the Division of Labor.* New York: Columbia University Press.

Madheswaran, S. and P. Attewell (2007). Caste Discrimination in the Indian Urban Labour Market: Evidence from the National Sample Survey. *Economic and Political Weekly 42(41), 4146-4153.*

Makinen, M., H. Waters, M. Rauch, N. Alamagambetova, R. Bitran, L. Gilson, S. Ram, G. Ubilla, D. Mclntyre and S. Pannarunothai (2000). Inequalities in health care use and expenditures: empirical data from eight developing countries and countries in transition. *Bulletin of the World Health Organization 78(1), 54-64.*

Matouschek, N. and I. Rasul (2008). The Economics of the Marriage Contract: Theories and Evidence. *Journal of Law and Economics 51, 59-110.*

Mensch, B. S., S. Singh and J. B. Casterline (2005). Trends in the Timing of First Marriage Among Men and Women in the Developing World. Website. `http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.175.7458&rep=rep1&type=pdf`

Mills, A. and L. Gilson (1988). Health Economics for Developing Countries: A survival Kit. Website. `http://r4d.dfid.gov.uk/PDF/Outputs/HealthEcFin_KP/WP01_`

88.pdf

Ministerial Leadership Initiative (2010). Reducing Financial Barriers to Reproductive Health Care Nepal Spotlight.

Website. http://www.ministerial-leadership.org/sites/default/files/events/event_files/Nepal

Ministry of Finance, Nepal (2011). Economic Survey of Nepal 2011. Website. http://www.mof.gov.np//en/archive-documents/economic-survey-21.html

Ministry of Health, Government of Nepal (2006). Nepal Health Sector Programme-Implementation Plan (NHSP-IP) 2004-2009. Website. http://un.org.np/node/10321

Ministry of Health and Population (2009). Nepal's National Health Account. Website. http://www.who.int/nha/country/npl/nepalnha.pdf

Moore, M. (2006). State Building in Nepal: Some Medium Term Policy Ideas. Website. http://www.bing.com/search?q=State+Building+in+Nepal%3A+Some+Medium+Term+Policy+Ideas&form=CPDTDF&pc=CPDTDF&src=IE-SearchBox

Munshi, K. and M. Rosenzweig (2006). Traditional Institutions Meet the Modern World: Caste, Gender and schooling choice in a globalizing economy. *American Economic Review 96(4), 1225-1252.*

Neugarten, B. L., J. W. Moore, and J. C. Lowe (1965). Age norms, age constraints, and adult socialization. *American Journal of Sociology 70(6), 710-717.*

Schooling Choice in a Globalizing Economy. *American Economic Review 96(4), 1225-1252.*

Oaxaca, R. (1973). Male-Female Wage Differentials in Urban Labor Markets. *International Economic Review 14(3), 693-709.*

Oppenheimer, V. K., M. Kalmijn and N. Lim (1997). Men's Career Development and Marriage Timing During a Period of Rising Inequality. *Demography 34(3), 311-*

330.

Pasqua, S. (2005). Gender Bias in Parental Investments in Children's Education: A Theoretical Analysis. *Review of Economics of the Household 3(3), 291-314.*

Patrinos, H. (1997). Differences in education and earnings across ethnic groups in Guatemala. *The Quarterly Review of Economics and Finance 37(4), 809-821.*

Phelps, E. (1972). The Statistical Theory of Racism and Sexism. *The American Economic Review 62(4), 659-661.*

Pudney, S. and M. Shields (2000). Gender, race, pay and promotion in the British nursing profession: estimation of a generalized ordered probit model. *Journal of Applied Econometrics 15 (4), 367-399.*

Riley, G. (2012). Discrimination in the labor market. Website. `http://www.tutor2u.net/economics/revision-notes/a2-micro-labour-market-discrimination.html`

Rosenbusch, N., R. Andreas, S. C. Parker and J.M. Unger (1994). Human Capital, Gender and Entrepreneurial Success: Empirical Evidence from China and Germany.

Rosenzweig, M. R. T. P. Schultz (1982). Market Opportunities, Genetic Endowments, and Intrafamily Resource Distribution: Child Survival in Rural India. *American Economic Review 72(4), 803-815.*

Roy, K. and D. H. Howard (2007). Equity in out-of-pocket payments for hospital care: evidence from India. *Health Policy 80(2), 297-307.*

Sawada, Y. (1997). Human Capital Investments in Pakistan: Implications of Micro Evidence from Rural Households. *The Pakistan Development Review 36(4), 695-712.*

Schmidt, C. M. and K. F. Zimmermann (1991). Work Characteristics, Firm Size and Wages. *The Review of Economics and Statistics 73(4), 705-710.*

Schmidt, P. J. and R. Strauss (1975). The Prediction of Occupation Using Multiple

Logit Model. *International Economic Review 16, 471-485.*

Schoen, C. and M. M. Doty (2004). Inequities in access to medical care in five countries: findings from the 2001 Commonwealth Fund International Health Policy Survey. *Health Policy 67, 309-322.*

Schultz, T. W. (1961). Investment in Human Capital. *The American Economic Review 51(1), 1-17.*

Singh, S. and R. Sharma (1996). Early Marriage among Women in Developing Countries. *International Family Planning Perspectives 22(4), 148-175.*

Subbarao, K., L. Raney and H. Dundar and J. Haworth (1994). Women in Higher Education: Progress, Constraints, and Promising Initiatives. Website. `http://publications.worldbank.org/index.php?main_page=product_info&cPath=0&products_id=20830`

Subedi, J. (1992). Primary Health Care and Medical Pluralism Exemplified in Nepal: A Proposal for Maximizing Health Care Benefit. *Sociology Focus 25(4), 321-328.*

Subramanian S.V., S. Nandy, M. Irving, D. Gordon, H. Lambert and G. D., Smith (2006). The Mortality Divide in India: The Differential Contributions of Gender, Caste, and Standard of Living Across the Life Course. *American Journal of Public Health 96(5), 818-825.*

Szelewicki, M. and T. Joanna (2009).Labour Market Racial Discrimination in South Africa Revisited. Website. `http://mpra.ub.uni-muenchen.de/16440/`

The World Bank (1988). Financing Health Services in Developing Countries. An Agenda for Reform. Website. `http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/1999/09/23/000178830_98101903343495/Rendered/PDF/multi_page.pdf`

UNICEF (2011). The State of World's Children 2011. Website. `http://www.unicef.org/sowc2011`

Van de Poel, E. and N. Speybroeck (2009) . Decomposing malnutrition inequalities between Scheduled Castes and Tribes and the remaining Indian population. *Ethnicity Health 14(3), 271-287.*

Van Der Heyden, J.H.A. and S. Demarest and J. Tafforeau and H. Van Oyen (2003) . Socio-economic differences in the utilisation of health services in Belgium. *Health Policy 65(2), 153-165.*

van Doorslaer, E. and A. Wagstaff (1992). Equity in the delivery of health care: Some international comparisons. *Journal of Health Economics 11, 389-411.*

van Doorslaer, E. and A. Wagstaff and H. van der Burga and T. Christiansen and D.D. Graeve and I. Duchesne and Ulf-G. Gerdtham and M. Gerfin and J. Geurts and L. Gross and U. Hakkinen and J. John and J. Klavus and R.E. Leu and B. Nolan and O. O'Donnell and C. Propper and F. Puffer and M. Schellhorn and G. Sundberg and O. Winkelhake (2000). Equity in the delivery of health care in Europe and the US. *Journal of Health Economics 19(5), 553-583.*

van Doorslaer, E. and C. Masseria and X. Koolman (2006). Inequalities in access to medical care by income in developed countries. *Canadian Medical Association Journal 174(2), 177-183.*

van Doorslaer, E. and X. Koolman (2004). Explaining the Differences in income-related health inequalities across European countries. *Health Economics 13(7), 609-628.*

Vietorisz, T. and B. Harrison (1973). Labor Market Segmentation: Positive Feedback and Divergent Development. *The American Economic Review 63(2), 366-376.*

Wagstaff, A. (2002). Poverty and health sector inequalities. *Bulletin of the World Health Organization 80(2), 97-105.*

Wagstaff, A. and E. van Doorsalaer (2000). Equity in Health Care Finance and Delivery. In *Handbook of Health Economics.* Amsterdam: Elsevier.

Wan, T. T. and L.C. Gray (1978). Differential Access to Preventive Services for Young Children in Low-Income Urban Areas. *Journal of Health and Social Behavior 19(3), 312-324.*

WHO (2007). Health System in Nepal: Challenges and Strategic Options. Website. http://www.nep.searo.who.int/LinkFiles/Health_Information_HSC.pdf

WHO (2009). WHO Global Health Expenditure Atlas. Website. http://www.who.int/nha/atlas.pdf

William, A. and R. Cookson (2000). Equity in Health. In *Handbook of Health Economics.* Amsterdam: Elsevier.

Winkleby, M. A., D. E. Jatulis, E. Frank and S. P. Fortmann (1992). Socioeconomic Status and Health: How Education, Income and Occupation Contribute to Risk Factors for Cardiovascular Disease. *American Journal of Public Health 82(6), 816-820.*

Woodhall, M. (1973). Investment in women: A reappraisal of the concept of human capital. *International Review of Education 19(1), 9-29.*

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data.* Cambridge: The MIT Press.

Wu, S., R. Wang, Y. Zhao, X. Ma, M. Wu, X. Yan and J. He (2013). The relationship between self-rated health and objective health status: a population-based study. *BMC Public Health 13(1), 313-320.*

Yang, D. T. and X. Zhu (2003). Household Production and Parental Investment in Children. Website.http://www.econ.yale.edu/conference/neudc03/papers/.pdf

Yokozeki, Y. (2000). Gender in Education and Development. *Journal of International Cooperation in Education 1(1), 45-63.*

Yueh, L. Y. (2001). A Model of Parental Investment on Children's Human Capital. Website. http://www.skope.ox.ac.uk/sites/default/files/SKOPEWP15.pdf

Appendix

Appendix 1: Variables definition (Chapter-1).

*Educ*1	Years of schooling (the highest level completed).
*Educ*2	Years of schooling imputed as 0 for those who did not report the highest level completed but reported as they never attended school.
Age:25-29	Taking value 1 if an individual's age was reported between 25-29 years; 0 otherwise.
Age:30-34	Taking value 1 if an individual's age was reported between 30-34 years; 0 otherwise.
Age:35-39	Taking value 1 if an individual's age was reported between 35-39 years; 0 otherwise.
Age:40-44	Taking value 1 if an individual's age was reported between 40-44 years; 0 otherwise.
Age:45-49	Taking value 1 if an individual's age was reported between 45-49years; 0 otherwise.
Urban	Taking value 1 if respondent was born in urban area; 0 Otherwise.
Mage	Age at marriage.
Feduc	Father's the highest level of education.
Meduc	Mother's the highest level of education.
Lnholding	Price of land-holding by a household.
Hincome	Household gross income calculated from farm-earning plus earning from sale of live-stock plus income from non-farm enterprises plus remittance received
Maithili	Taking value 1 if an individual's language is Maithili; 0 otherwise.

Appendix 1: Continued.

Avmage	Average age at marriage derived from the interaction between ethnicity and region.
Brahman	Taking value 1 if a respondent's ethnicity was reported as Brahman; 0 otherwise.
Chhetri	Taking value 1 if a respondent's ethnicity was reported as Chhetri; 0 otherwise.
Newar	Taking value 1 if a respondent's ethnicity was reported as Newar; 0 otherwise.
Magar	Taking value 1 if a respondent's ethnicity was reported as Magar; 0 otherwise.
Tharu	Taking value 1 if a respondent's ethnicity was reported as Tharu; 0 otherwise.
Tamang	Taking value 1 if a respondent's ethnicity was reported as Tamang; 0 otherwise.
Kami	Taking value 1 if a respondent's ethnicity was reported as Kami; 0 otherwise.
Yadav	Taking value 1 if a respondent's ethnicity was reported as Yadav; 0 otherwise.
Muslim	Taking value 1 if a respondent's ethnicity was reported as Muslim; 0 otherwise.
Rai	Taking value 1 if a respondent's ethnicity was reported as Rai; 0 otherwise.
Gurung	Taking value 1 if a respondent's ethnicity was reported as Gurung; 0 otherwise.
Limbu	Taking value 1 if a respondent's ethnicity was reported as Limbu; 0 otherwise.
Sarki	Taking value 1 if a respondent's ethnicity was reported as Sarki; 0 otherwise.
Other	Taking value 1 if a respondent's ethnicity was reported as Other; 0 otherwise.
Urban	Taking value 1 if individual was born in urban areas; 0 otherwise.
Eastern	Taking value 1 if an individual was born in eastern development region; 0 otherwise.
Central	Taking value 1 if an individual was born in central development region; 0 otherwise.
Western	Taking value 1 if an individual was born in western development region; 0 otherwise.
Mid-western	Taking value 1 if an individual was born in mid-western development region; 0 otherwise.
Far-western	Taking value 1 if an individual was born in far-western development region; 0 otherwise.
Abroad	Taking value 1 if an individual was born in abroad; 0 otherwise.

Appendix 2: Variables definition (Chapter-2).

Tagadhari	Taking value 1 if an individual's ethnicity is reported as Brahman, Chhetri, Newar and Yadav; 0 otherwise.
Matwali	Taking value 1 if an individual's ethnicity is reported as Gurung, Magar, Tharu, Tamang, Rai and Limbu; 0 otherwise.
Pani Nachalne	Taking value 1, if an individual's ethnicity is reported as Damai, Kami, Sarki and Muslim; 0 otherwise.
Lhwage	log of hourly wage (cash, in-kind, bonus, transport, and medical allowances).
Education	Years of schooling completed (the highest level completed).
Experience	Age-years of schooling-6.
Married	Taking value 1 if an individual was married; 0 otherwise.
Lnholding	Price of landholding by a household.
Rural	Taking value 1 if an individual was born in rural area; 0 otherwise.
Small firm	Taking value 1 if a firm employs only one employee; 0 otherwise.
Medium firm	Taking value 1 if a firm employs $2-10$ employees; 0 otherwise.
Large firm	Taking value 1 if a firm employs more than 10 employees; 0 otherwise.
Eastern	Taking value 1, if an individual lives in eastern administrative region; 0 otherwise.
Central	Taking value 1 if an individual lives in central administrative region; 0 otherwise.
Western	Taking value 1 if an individual lives in eastern administrative region; 0 otherwise.
Mid-western	Taking value 1 if an individual lives in mid-western administrative region; 0 otherwise.
Far-western	Taking value 1 if an individual lives in far-western administrative region; 0 otherwise.
Abroad	Taking value 1 if an individual lives outside Nepal; 0 otherwise.
Unskilled	Taking value 1 if an individual's occupation is not included in other categories; 0 otherwise.
Professional	Taking value 1 if an individual's occupation was reported as doctor, engineer, administrative executive, religious professional; 0 otherwise.

Appendix 2: Continued.

Clerical	Taking value 1 if an individual's occupation was reported as clerk, typist, book keeper, telephone operator, military, other clerical; 0 otherwise.
Service	Taking value 1 if an individual's occupation was reported as travel, trekking, cooking, housekeeping, care takers, laundry workers, barbers and other service worker; 0 otherwise.
Sales	Taking value 1 if an individual's occupation was reported as shop and stall sales person; 0 otherwise.
Agri-worker	Taking value 1 an individual's occupation was reported as farm manager, farm worker, agricultural worker, forestry worker, fisherman, hunters and trapper; 0 otherwise.
Skilled	Taking value 1 if an individual's occupation was reported as metal processor, chemical processor, plumber, welders, jewelery workers, paper makers; 0 otherwise.
Agricultural	Taking value 1 if the type of industry was reported as agricultural, forestry and logging and fishing; 0 otherwise.
Mining	Taking value 1 if the type of industry was reported as coal mining, petroleum gas, metal mining and other mining; 0 otherwise.
Manufacturing	Taking value 1 if the type of industry was reported as food and beverage, textile apparel, wood furniture, paper printing, handicrafts, other metallic; 0 otherwise.
Construction	Taking value 1 if the type of industry was reported buildings, street highways, water ports project, irrigation, electricity gas and water; 0 otherwise.
Trade	Taking value 1 if the type of industry was reported as wholesale, retail and restaurant; 0 otherwise.
FRE	Taking value 1 if the type of industry was reported as finance, insurance and real estate; 0 otherwise.
Servicesec	Taking value 1 if the type of industry was reported as transport, communication, recreation and cultural and international; 0 otherwise.
Others	Taking value 1 if the type of industry is not included in above categories; 0 otherwise.

Appendix 3.A: Variables definition

Log OOP	Log of OOP expenditure which includes consultation fees, travel and medicine costs.
Lincome	Log of per capita consumption (food and frequently purchased non-food expenditure).
Lnholding	Price of landholdings by a household.
Rural	Taking value 1 if an individual resides in rural areas; 0 otherwise.
Male	Taking value 1 if an individual is male; 0 otherwise.
Hhsize	Number of members in a household.
Lhdi	Taking value 1 if individual belongs to 22 districts with the lowest HDI; 0 otherwise.
Age:0-34	Taking value 1 if an individuals' age was reported within 0-34 years; 0 otherwise.
Age:35-44	Taking value 1 if an individual's age was reported within 35-44 years; 0 otherwise.
Age:45-54	Taking value 1 if an individual's age was reported within 45-54 years; 0 otherwise.
Age:55-64	Taking value 1 if an individual's age was reported within 55-64 years; 0 otherwise.
Age:65-74	Taking value 1 if an individual's age was reported within 65-74 years; 0 otherwise.
Age$\geq$ 75	Taking value 1 if an individual's age was reported $>$ 74 years; 0 otherwise.

Appendix 3.A: Continued

Illiterate	Taking value 1 if an individual's level of education was reported as 0; 0 otherwise.
Primary	Taking value 1 if an individual's educational level was reported within 1-5 years of schooling; 0 otherwise.
High school	Taking value 1 if an individual's educational level was reported within 6-10 years of schooling; 0 otherwise.
University	Taking value 1 if an individual's educational level was reported >10 years of schooling; 0 otherwise.
Far-western	Taking value 1 if an individual was living in far-western region; 0 otherwise.
Central	Taking value 1 if an individual was living in central region; 0 otherwise.
Western	Taking value 1 if an individual was living in western region; 0 otherwise.
Mid-western	Taking value 1 if an individual was living in mid-western region; 0 otherwise.
Abroad	Taking value 1 if an individual was living in abroad; 0 otherwise.
Eastern	Taking value 1 if an individual was living in eastern region; 0 otherwise.
Chronic illness	Taking value 1 if an individual was reported to have chronic illness; 0 otherwise.
Other illness	Taking value 1 if an individual was reported to have illness other than chronic illness; 0 otherwise.
Not ill	Taking value 1 if an individual was reported not to have any illness; 0 otherwise.

Appendix 3.A: Continued

Illness started<= One year	Taking value 1 if an individual was reported to have chronic illness from a year or less; 0 otherwise.
Illness started<= five years	Taking value 1 if an individual was reported to have chronic illness from more than a year but less than five years; 0 otherwise.
Illness started<= ten years	Taking value 1 if an individual was reported to have chronic illness from more than five years but less than ten years; 0 otherwise.
Illness started>= ten years	Taking value 1 if an individual was reported to have chronic illness from ten and more years; 0 otherwise.
Work missed≤week	Taking value 1 if work missed due to chronic illness was reported as $\leq$ 7 days; 0 otherwise.
Work missed≤month	Taking value 1 if work missed due chronic illness was reported as 8-30 days; 0 otherwise.
Work missed>month	Taking value 1 if work missed due to chronic illness was reported as $>$ 30 days; 0 otherwise.
Excellent	Taking value 1 if self-assessed health status was reported as excellent; 0 otherwise.
Good	Taking value 1 if self-assessed health status was reported as good; 0 otherwise.
Fair	Taking value 1 if self-assessed health status was reported as fair; 0 otherwise.
Poor	Taking value 1 if self-assessed health status was reported as poor; 0 otherwise.
Very poor	Taking value 1 if self-assessed health status was reported as very poor; 0 otherwise.

Appendix 3.B: Human development indicator by districts in Nepal.

HDI	Name of districts
HDI $\leq$ 0.2	Kalikot, Bajura and Mugu.
0.2< HDI $\leq$ 0.3	Kailali, Dang, Sindhuli, Darchula, Kapilbastu, Sindhupalchok, Rukum, Rolpa, Dadeldhura, Dhading, Baitadi, Dailekh, Rasuwa, Humla, Achham, Jumla, Salyan, Doti, Dolpa, Jajarkot and Bajhang.
0.3<HDI $\leq$ 0.4	Bhaktapur, Terhathum, Tanahun, Sunasari, Ilam, Kabhrepalanchok Syangja, Lamjung, Saptari, Chotwan, Sankhuwasabha Taplejug, Rupndehi, Parwat, Surkhet, Parsa, Udayapur, Solukhumbu Bhojpur, Siraha, Okhaldhunga, Dolakha, Baglung, Palpa, Kanchanpur Arghakhanchi, Dhanusha, Panchthar, Sarlahi, Gulmi, Pyuthan Mahottari, Khotang, Mustang, Ramechhap, Nuwakot, Myagdi, Bara Makwanpur, Rautahat, Banke, Gorkha, Manang, Bardiya and Nawalparasi.
0.4< HDI $\leq$0.5	Kaski, Morang, Jhapa and Dhankuta.
0.5< HDI $\leq$0.6	Lalitpur.
HDI > 0.6	Kathmandu.

Source: Nepal Human Development Report, 1998.

Note: HDI is a tool developed by the United Nation that ranks the level of social and economic development of a specific region based on four criteria: life expectancy at birth, average years of schooling, expected years of schooling and gross national income per capita.