The Impact of Migrant Remittances on Rural Labor Supply: Evidence from Nepal

Amanda Wong

Abstract: Migrant remittances are one of the largest forms of international financial flows, exceeding the total flows of official development assistance, and second only to foreign direct investment. Consequently, remittances have played an increasingly important role in the economic development of many countries. In Nepal, remittances have been shown to contribute to significant levels of poverty reduction, increased private consumption, and reduced credit constraints. However, the question of whether migrant remittances increase or decrease the domestic labor supply of remaining household members is not well understood in the economics literature due to the competing nature of the income and substitution effects. Using data from the 2016 Nepal Household Risk and Vulnerability Survey, this paper uses a variety of empirical strategies to estimate the impact of migrant remittances on domestic rural labor supply. Firstly, this paper uses an instrumental variables approach, with the historic share of international migrants at a district level as an instrument for remittances. This paper then leverages the large number of control characteristics available in the survey to estimate this relationship through a propensity score matching method, a random forest regression, and a double lasso variable selection regression model. This paper finds that, through the ordinary least squares method, propensity score matching method, and post double lasso variable selection method, belonging to a remittance household reduces an individual's daily time worked by 6 minutes whereas, through an instrumental variables approach, this effect is much larger at a reduction of around 3 hours. While the magnitude of the effects of remittances on daily labor supply differs by estimation strategy, this paper finds a consistently negative relationship between remittances and rural labor supply.

[1] Acknowledgements: I would like to extend my gratitude to Professor Frederico Finan for advising me throughout the thesis writing process. I would also like to thank Professor Keiko Yamanaka, Nicholas Swanson, and the Berkeley Economic Review Peer Review team for all their insights and feedback. All errors and shortcomings are my own.

Table of Contents

I. Introduction

Remittances are defined as a financial inflow arising from the cross border movement of nationals of a country sent by migrant workers to their country of origin[2]. In recent years, remittances have become one of the largest sources of monetary flows to low and middle income countries, with amounts exceeding official development assistance and second only to foreign investment[3]. Consequently, remittances have profound implications for the development of many emerging economies. Remittances have been shown to reduce credit constraints, increase private consumption, and lift millions of households out of poverty[4]. However, remittances have also resulted in increased dependence on foreign countries, and instances of brain drain related to an outmigration of skilled residents[5].

The quantity and quality of a nation's labor supply is one of the key determinants of its long term economic growth[6]. However, the impact of migrant remittances on local labor supply of remaining household members of migrants remains unclear. On one hand, remittances may negatively impact the labor supply of left behind family members through the income effect. As remittances provide an additional source of income to households, this raises the hourly reservation wages of family members of migrants, inducing workers to reduce their hours of labor supplied to the market. On the other hand, remittances may increase the number of hours worked by remaining household members through the substitution effect. This is because a reduction in the pool of workers due to migration abroad raises the average local market wage

---

[2] "Remittances." *Migration Data Portal*, migrationdataportal.org/themes/remittances.
[3] Murakami, Enerelt, et al. "The Impact of Migration and Remittances on Labor Supply in Tajikistan." *Journal of Asian Economics*, vol. 73, 2021, p. 101268., https://doi.org/10.1016/j.asieco.2020.101268.
[4] Migration in Nepal, A Country Profile. International Organization for Migration. 2019.
[5] Murakami, Enerelt, et al. "The Impact of Migration and Remittances on Labor Supply in Tajikistan." *Journal of Asian Economics*, vol. 73, 2021, p. 101268., https://doi.org/10.1016/j.asieco.2020.101268.
[6] Chami, Ralph, et al. "Are Remittances Good for Labor Markets in Lics, Mics and Fragile States?" *SSRN Electronic Journal*, 2018, https://doi.org/10.2139/ssrn.3221121.

for a variety of jobs, which may increase the average number of hours supplied by workers. Other theories as to why remittances may increase the local labor supply of left behind family members include the idea that remittances may lift liquidity constraints and create more opportunities for household members to engage in entrepreneurship activities. Another theory is that left behind household members of migrants, particularly in agriculture, may increase their local labor supply in order to replace the work that would have been supplied by the household member who migrated abroad. Consequently, the link between remittances and labor supply is unclear, and its impact may differ on a country by country basis.

Nepal is one of the poorest nations in the world, with 25.2% of its population earning less than $1.25 a day[7]. The predominant sector in Nepal is agriculture, employing two thirds of the nation's population, and 79.74% of its population living in rural areas[8]. However, despite being a largely agrarian country, an increase in the inflow of migrant remittances in recent decades have fueled its economic growth and development. Currently, Nepal ranks as the sixth top remittance receiving country in the world by percentage of GDP (24.1%) [9]. Remittances have fueled most of Nepal's economic growth in recent decades, as Nepal living standards surveys have shown that remittance income has significantly contributed to reducing its headcount poverty rate from 42 percent in 1995 to 31 percent in 2003 to 25 percent in 2010[10]. Remittances have profound implications for Nepal's economic development, as they have been shown to increase women's educational attainment, fuel Nepal's rapid GDP growth, increase

[7] Phadera, Lokendra. "Impact of International Migration on Labor Supply in Nepal." 2019, https://doi.org/10.1596/1813-9450-9014.
[8] "Nepal." *IFAD*, International Fund for Agricultural Development, https://www.ifad.org/en/web/operations/w/country/nepal.
[9] "Personal Remittances, Received (% of GDP)." *The World Bank* , 2021, https://data.worldbank.org/indicator/BX.TRF.PWKR.DT.GD.ZS?most_recent_value_desc=true.
[10] Thapa, Sridhar; Acharya, Sanjaya (2017) : Remittances and household expenditure in Nepal: Evidence from cross-section data, Economies, ISSN 2227-7099, MDPI, Basel, Vol. 5, Iss. 2, pp. 1-23, http://dx.doi.org/10.3390/economies5020016

consumption and investment in human capital, increase nutrition, increase wages, and significantly reduce poverty rates[11].

However, although remittances have undoubtedly contributed significant economic and welfare gains, one potential concern of Nepal's high reliance on remittance income is its impact on the economic activity within its borders, particularly the labor supply of the family members of remittance workers left behind. Remittances have created a large absentee population in Nepal. According to Nepal's Central Bureau of Statistics, 55.8% of households sampled in Nepal in 2010 received remittances[12], and out of the total population aged 5 years and above, 37 percent have migrated to other places.[13]

This paper seeks to contribute an insight into the role of remittances on the labor supply of Nepal's remittance receiving households in non-metropolitan settings. The primary dataset used in this paper is the 2016 Nepal Household Risk and Vulnerability Survey from the World Bank, which is a multi-topic survey conducted across 6,000 households in the non-metropolitan areas of Nepal[14]. In order to estimate the causal impact of remittances on rural labor supply of remaining household members in Nepal, I employ multiple econometric techniques. The first econometric technique I use is an instrumental variables approach, in which I use the share of international migrants in 2010 on the district level as an instrument for remittances in 2016. In order to overcome some potential limitations of my instrument and take advantage of the wealth of control characteristics from the survey, I also use a propensity score matching method, a random forest propensity score method, and a post double lasso variable selection regression to

[11] Phadera, Lokendra. "Impact of International Migration on Labor Supply in Nepal." *World Bank Group - Poverty and Equity Global Practice*, Sept. 2019, doi:10.1596/1813-9450-9014.
[12] Thapa, Sridhar, and Sanjaya Acharya. "Remittances and Household Expenditure in Nepal: Evidence from Cross-Section Data." *Economies* 5.2 (2017): 16. *Crossref*. Web.
[13] Nepal Living Standards Survey, 2010/2011 Highlights. Central Bureau of Statistics.
[14] Walker, Thomas and Jacoby, Hanan. *Household Risk and Vulnerability Survey 2016*. The World Bank. 17 September 2021.

estimate the relationship between household remittances and rural labor supply in the Nepalese context.

Background of Nepalese Migration

In recent decades, massive rates of outmigration has transformed Nepal both economically and culturally. Currently, more than half of all Nepali households have a family member who has migrated for work, and an estimated 2 million people are working abroad, constituting 8.3% of Nepal's total population[15]. Of the total absentee population, 87.6 percent of workers are male, and 12.4 percent are female[16]. Consequently, the absence of young-working age men in Nepal has had significant implications on the role of women in these left behind households. Firstly, women have been reported to fill niches in the labor market and in the household that were previously filled by men. For example, instead of being contributing farm laborers, with the absence of men, women may become primary farmers. Moreover, remittances have also been shown to increase self-employment of women, increase women's educational attainment, and increase the role of women as heads of households[17].

Foreign employment is the main motivation for international migration, with 8 out of 10 migrants migrating for work-related reasons, and the majority of the permits are for purposes of low-skilled construction work, where migrants engage in work described as "triple D" or "difficult, dangerous, and dirty"[18]. Currently, the top destination countries for Nepalese workers include India, Malaysia, Saudi Arabia, Qatar, United Arab Emirates, and Kuwait. According to

---

[15] *Nepal Labor Market Update* . International Labor Organization, Jan. 2017.
[16] Migration in Nepal, A Country Profile. International Organization for Migration. 2019.
[17] Slavchevska, Vanya, et al. "Rural Outmigration and the Gendered Patterns of Agricultural Labor in Nepal." 2020, doi:10.2499/p15738coll2.134190.
[18] Ghimire, Dirgha, and Bhandari, Prem. "Study of Migration and Later Life Health in Nepal." *Journal of Migration and Health*, vol. 1-2, 2020, p. 100018., https://doi.org/10.1016/j.jmh.2020.100018.

the Ministry of Finance, Nepalese households received a total of 430 billion and 560 billion rupees in the years of 2012 and 2014[19].

A combination of economic, environmental, political, and social factors have motivated many Nepali citizens to migrate abroad. Firstly, a lack of decent work opportunities at home, coupled with higher earning opportunities abroad, has caused many Nepalis, especially young men, to work abroad. Given that agriculture employs 60% of Nepal's labor force, an increase in environmental degradation and natural disasters such as deforestation, recurrent floods, and landslides, has also led to many households to be exposed to increased levels of risk, causing Nepalis to use migration as a risk diversification strategy against poor agricultural conditions[20].

A number of treaties signed by the government of Nepal has prompted much of this migration. Since the establishment of the 1950s Treaty of Peace and Friendship, Nepal has shared an open border with India, and thus migrants are able to cross the border without any documentation requirements or limits. Additionally, the Employment Act of 2007 signed by Nepal was designed to provide security, protect the welfare of migrants, provide migrants with education and training before leaving the country, and monitor businesses that facilitate migration processes. In the last decade, Nepal has also signed international labor treaties with Qatar, the United Arab Emirates, the republic of Korea, Bahrain, Japan, and Malaysia[21].

Lastly, social factors are also important in driving the increase in migrants in recent years. The number of family and community members that an individual migrant has in the destination country increases their propensity to migrate abroad. For example, in the Nepalese

[19] Lokshin, Michael, and Elena Glinskaya. "The Effect Of Male Migration For Work On Employment Patterns Of Females In Nepal." *Policy Research Working Papers*, 2008, doi:10.1596/1813-9450-4757.
[20] Migration in Nepal, A Country Profile. International Organization for Migration. 2019.
[21] *Nepal Labor Market Update* . International Labor Organization, Jan. 2017.

village of Malma, 1800 individuals out of an entire population of 6,400 currently work abroad in restaurants in Japan[22]. Studies have also shown that migration is also based on ethnicity, as Muslim workers have been reported to be more likely to migrate to Gulf countries, whereas Hill Dalit workers have an increased propensity to migrate to India[23]. Consequently, migrant networks are another significant factor in an individual migrant's decision to migrate.

Migration has undoubtedly deeply transformed Nepal both socially and culturally. Migration has become viewed as a "right of passage" for many young Nepalese men. However, whilst remittances have contributed positively to Nepal's economic growth, many migrants face significant risks and challenges once in their destination country. Nepali migrants have been reported to face difficult working conditions and exposed to high levels of harassment and abuse such as untimely payments and passport confiscations. Additionally, high death rates of young Nepali migrant workers have been reported due to abuse by employers[24].

Interestingly, the relationship between wealth and migration is not necessarily clear. Migration involves upfront costs for travel, and thus the poorest parts of the population may not be likely to migrate as they are unable to afford the upfront costs of migration. On the other hand, while wealthier households may be able to afford the upfront costs, but they may not have the same motivations to do so as they are not as enticed by higher earnings abroad, nor face the same push factors that less wealthy households do[25].

---

[22] Kharel, Dipesh, From *Lahures* to Global Cooks: Network Migration from the Western Hills of Nepal to Japan , *Social Science Japan Journal*, Volume 19, Issue 2, Summer 2016, Pages 173–192, https://doi.org/10.1093/ssjj/jyw033

[23] Gurung, Yogendra Bahadur (2012) "Migration From Rural Nepal: A Social Exclusion Framework," Himalaya, the Journal of the Association for Nepal and Himalayan Studies: Vol. 31: No. 1, Article 12. Available at: http://digitalcommons.macalester.edu/himalaya/vol31/iss1/12

[24] *Nepal Labor Market Update* . International Labor Organization, Jan. 2017.

[25] Gurung, Yogendra Bahadur (2012) "Migration From Rural Nepal: A Social Exclusion Framework," Himalaya, the Journal of the Association for Nepal and Himalayan Studies: Vol. 31: No. 1, Article 12. Available at: http://digitalcommons.macalester.edu/himalaya/vol31/iss1/12

Background of Nepalese Agricultural Labor Market

Nepal's labor market is largely dominated by the agricultural sector, as it employs 60% of the total labor force. Much of Nepal's agricultural sector is subsistence in nature, with wage employment only constituting 16.9% of the labor force[26]. Rice, maize, millet, wheat, and buckwheat are major staple food crops, as well as oilseeds, potato, tobacco, sugarcane, jute and cotton[27]. Nepal's agricultural land is classified into three types: Terai (grasslands), hill, and mountain regions. The Terai regions are known as the most agricultural productive regions, producing 64% of the nation's agriculture, 42% of the middle mountain's land is used for agriculture, and 13% of the area in the high mountains is used for agriculture[28]. In general, productivity in the agricultural sector is low as it only contributes to 30% of GDP, and farming still relies primarily on manual labor and rainfall[29]. The lack of modernized farming methods and difficulties with irrigation and transport also implies that Nepal does not have the capacity to support the local population with current food production levels. The incidence of poverty is most severe in the mountains, with 56% of the population in the mountain region classified as poor and ultra poor, and face frequent food shortages.

Analyzing the impact of remittances on labor supply, especially in the agricultural setting, is an interesting question. In Nepal, similar to other countries in the Asia Pacific region, poverty is most highly concentrated in rural areas. Most of these rural households are engaged in agriculture and are more vulnerable to severe and frequent shocks compared to urban households. As a result, many rural households tend to resort to alternative strategies to cope

[26] *Nepal Labor Market Update* . International Labor Organization, Jan. 2017.
[27] *Nepal at a Glance* . Food and Agriculture Organization of the United Nations, 2021, https://www.fao.org/nepal/fao-in-nepal/nepal-at-a-glance/en/ .
[28] "Nepal." *IFAD*, International Fund for Agricultural Development, https://www.ifad.org/en/web/operations/w/country/nepal.
[29] *Nepal Labor Market Update* . International Labor Organization, Jan. 2017.

with these temporary shocks, such as switching to nonfarm employment or business creation. Migration -- permanent, temporary, seasonal, or cyclical -- is another common coping strategy among rural households and is often viewed as a tool to counter seasonal poverty, especially if nonfarm employment or business creation are infeasible solutions in the area[30]. Consequently, examining how remittances impact the labor supply, specifically in the agricultural setting, is a more relevant question as it has significant welfare implications. Firstly, because the rural agricultural population is amongst the poorest and most vulnerable population in Nepal, and secondly, because individuals in the agricultural sector are more likely to be impacted by remittances.

II. Literature Review

This paper adds to a variety of recent studies examining the relationship between labor markets and remittances in the case of Nepal. In a recent World Bank Working Policy Paper, Phadera (2019) examines the relationship between remittances and labor supply in Nepal by computing the percentage of international migrants from a village as an instrument for migration in 2010 to 2011. He finds that, on the intensive margin, adults from migrant-families decrease their weekly hours of labor supply for wage-employment by about eight hours when compared to the adults from non-migrant households[31]. Additionally, using data from the 2008 Nepal Living Standards survey, Sharma (2020) uses a similar instrumental variables approach to Phadera and finds that while remittances increase labor supply across the extensive margin in urban areas for females, there is a fall in the labor force participation rate in the agricultural

[30] Mobarak, Ahmed Mushfiq, and Maira Emy Reimão. "Seasonal Poverty and Seasonal Migration in Asia." *Asian Development Review*, vol. 37, no. 1, 2020, pp. 1–42., https://doi.org/10.1162/adev_a_00139.
[31] Phadera, Lokendra. "Impact of International Migration on Labor Supply in Nepal." *World Bank Group - Poverty and Equity Global Practice*, Sept. 2019, doi:10.1596/1813-9450-9014.

sector[32]. Sharma explains this fall in the labor force participation rate in the agricultural sector by stating that in Nepal, the agricultural sector is less attractive due to its low productivity with farm wages and that agricultural self employment hardly generates sufficient income for the household to achieve a high standard of living. Sharma's findings have implications for this paper, as it implies that households that have alternative sources of income in the form of remittances, farm work is usually less preferred, therefore that workers in the agricultural sector would be likely to reduce their hours greater than non-agricultural workers.

In the case of the agricultural sector in the Nepalese setting specifically, Slavchevska et al (2020) uses a multi-topic household survey from Nepal in 2017 administered by the FAO-World Bank in five districts. The authors use an instrumental variables approach to analyze the labor market decisions of men and women in the agricultural sector, using historical weather variability as an instrument. The authors find that when a family member migrates, women take on additional responsibilities on the farm and are 28.6 percentage points more likely to be self reemployed in agriculture compared to women without a family member who is a migrant. They also find that men in rural areas in general reduce their labor supply to any activity by nearly seven percentage points on the extensive margin.[33] Moreover Lokshin & Glinskaya (2004) using data from the 2004 of the National Living Standard Survey to find the impact of male migration on women's labor supply. In an approach similar to Phadera (2019), using historical migration rates from the 2001 Nepal Census Data of each surveyed ward as an

[32] Sharma, Hari, 2020. "The effect of emigration and remittances on labour supply of the left-behind: Evidence from Nepal," MPRA Paper 102091, University Library of Munich, Germany.
[33]Slavchevska, Vanya, et al. "Rural Outmigration and the Gendered Patterns of Agricultural Labor in Nepal." 2020, doi:10.2499/p15738coll2.134190.

instrumental variable, they find that having a male remittance worker decreases women's rates of market work participation by 5.4 percentage points.[34]

Propensity score matching methods have also been used to evaluate the impact of remittances in the Nepalese setting. Thapa et al (2017) used a propensity score matching method to evaluate the relationship between remittances and household expenditure in Nepal. Using the 2010/2011 National Living Standards Survey, they find that remittance income is more likely to increase the share of household budget allocated to the consumption of nonfood and investment in education and health, and leads to a decrease in the share of food consumption and other goods[35]. In the context of remittances, Siddiqui (2013) has also used a propensity score matching method to evaluate the impact of remittances on expenditure, and found that remittances have a positive impact on income, expenditure, savings, and human and physical capital accumulation in Pakistan[36]. However, both of these propensity score matching methods have been used to examine the impact of remittances on household consumption, and there have been no papers that have specifically used these methods to examine the impact on labor supply.

To my knowledge, there have been no papers in the remittance literature, and especially not in the Nepalese setting, to have used machine learning techniques to evaluate the impact of remittances on rural labor supply. There has been a growing interest in using machine learning methods for both propensity score estimation as well as control variable selection. A recent paper by Goller et. al. (2019) has argued that estimating propensity scores in matching type

---

[34] Lokshin, Michael, and Elena Glinskaya. "The Effect Of Male Migration For Work On Employment Patterns Of Females In Nepal." *Policy Research Working Papers*, 2008, doi:10.1596/1813-9450-4757.
[35]Thapa, Sridhar, and Sanjaya Acharya. "Remittances and Household Expenditure in Nepal: Evidence from Cross-Section Data." *Economies* 5.2 (2017): 16. *Crossref*. Web.
[36] Siddiqui, Rizwana. "Impact Evaluation of Remittances for Pakistan: Propensity Score Matching Approach." *The Pakistan Development Review*, vol. 52, no. 1, 2013, pp. 17–44., https://doi.org/10.30541/v52i1pp.17-44.

estimators with machine learning can help in variable selection, allow functional flexibility, and increase the precision of estimates by avoiding overfitting of the propensity score[37]. A random forest regression method can be used for propensity score estimation, and allow for a more accurate method to predict treatment participation, given the covariates, by trading off bias and variance in out of sample comparisons. By examining the ability of several measures of covariate balance in predicting the quality of propensity scores in terms of bias reduction, Massimo and Cannas (2019) found that random forests performed the best when propensity scores were used for matching[38], and thus I employ this technique in my analysis as an alternative method to estimate a propensity score.

Another machine learning technique that I use in this paper, and one that has not been used previously in the context of remittances nor the Nepalese setting, is a post-double lasso variable selection method. Urminsky et al (2016) has argued that using a double lasso variable regression method can be used to not over-select potentially spurious covariates, and through a series of simulations, have demonstrated that this method reduces Type I error and increases statistical power[39]. Additionally, Belloni et al (2014) has found that a post double selection estimation method is appealing for estimating parameters of high dimensional data such as surveys and censuses, where proportion of the numbers of variables to the number of observations[40]. By including covariates that predict both the outcome variable and the treatment variable identified by lasso regression, a post double variable selection method is able to reduce

---

[37] Goller, Daniel; Lechner, Michael; Moczall, Andreas; Wolff, Joachim (2019) : Does the Estimation of the Propensity Score by Machine Learning Improve Matching Estimation? The Case of Germany's Programmes for Long Term Unemployed, IZA Discussion Papers, No. 12526, Institute of Labor Economics (IZA), Bonn

[38] Cannas, Massimo, and Bruno Arpino. "A Comparison of Machine Learning Algorithms and Covariate Balance Measures for Propensity Score Matching and Weighting." *Biometrical Journal*, 2019, https://doi.org/10.1002/bimj.201800132.

[39] Urminsky, Oleg, et al. "Using Double-Lasso Regression for Principled Variable Selection." *SSRN Electronic Journal*, 2016, https://doi.org/10.2139/ssrn.2733374.

[40] Alexandre Belloni, Victor Chernozhukov, Christian Hansen, Inference on Treatment Effects after Selection among High-Dimensional Controls, *The Review of Economic Studies*, Volume 81, Issue 2, April 2014, Pages 608–650, https://doi.org/10.1093/restud/rdt044

omitted variables bias and allow for a more causal interpretation of the estimate. Consequently, I take advantage of the large number of variables and control characteristics in my dataset and use a post-double lasso variable estimation strategy in order to obtain a more causal estimate.

Whilst many past papers examining the impacts of remittances on rural labor markets in Nepal largely involve an instrumental variables strategy, none of these papers have specifically examined the impact of migrant remittances on labor supply across the intensive margin specifically in the agricultural setting. Consequently, I hope to contribute to the literature by using a similar instrumental variable of historic migrant shares by the district level to examine the relationship between remittances and rural labor supply. In addition, I hope to contribute to the literature by examining the relationship between remittances and rural labor supply by using experimental techniques that have not been previously used in this setting, including propensity score matching, random forest regression models, and a post-double lasso variable selection model. I then compare my estimates of the average treatment effect across these different estimation strategies.

III. Data and Descriptive Statistics

2016 Nepal Household Risk and Vulnerability Survey

The main data set that I use is the 2016 Nepal Household Risk and Vulnerability Survey from the World Bank. This is a cross sectional data set which is intended to provide the government of Nepal information on the patterns of exposure to shocks at the household level and on the vulnerability of households' welfare to these shocks. The survey covers 6,000 households in the non-metropolitan areas based on the 2010 census of Nepal. Since the survey only covers the non-metropolitan areas of Nepal, the data set includes households from 50 out of the 75 districts in Nepal. The household questionnaire covered 16 modules including the household roster, education, health, housing and access to facilities, food expenses and home production, non-food expenditures and inventory of durable goods, jobs and time use, wage jobs, farming and livestock, non-agriculture enterprises/activities, migration, credit, savings and financial assets, private assistance public assistance, shocks, and anthropometrics.

The section of the survey that I used to analyze the labor market conditions in rural Nepal is the jobs and time use section of the dataset, which was only limited to individual workers who are currently in the labor force. This section asks individual workers the number of hours that they worked on a given day, and whether or not they worked in each month in the Nepalese calendar, which allows me to study Nepal's labor supply along the intensive margin and labor participation for each month, but not the labor force participation on labor supply across extensive margin. I then merged this dataset with the migration portion of the dataset that asks whether or not the household received remittances in the past 12 months to identify remittance-receiving individual workers, as well as information regarding education, agriculture, savings, and consumption as controls. After the merging of all of the different portions of the survey, my final dataset included 7,722 observations of individual workers across 50 non-metropolitan districts in Nepal.

There were some inconsistencies regarding the cleaning of my dataset. Firstly, in terms of hours worked per day, there were a number of individuals that responded working a total of 24 hours a day, which is impossible. This may have been because a very large number of individuals in the dataset work in the agricultural sector and thus do not distinguish work from leisure. Another possibility may be due to measurement error in the way surveys were conducted, such as a difficulty in translation. Consequently, I dropped individuals that reported working more than 15 hours from my dataset, assuming that the average individual sleeps 9 hours, making 15 the theoretical maximum number of working hours. Another source of data concern would be that many individuals did not respond to the question of earnings per month from primary jobs. This may be due to the overrepresentation of agricultural workers in my dataset, who rely on subsistence farming for livelihoods, and may not actually receive wages from an employer. This explanation can be supported by the fact that agriculture employs 60% of Nepal's labor force and 85% of those workers rely on subsistence farming[41]. Consequently, I used a combination of the reported wages by wage workers and total value of goods received in kind by households, which is common amongst agricultural areas, as a proxy for wages. I added up the total wages reported with the total value received in kind to estimate the total compensation variable. However, due to the substinent nature of farming in Nepal, this total compensation variable is still not a perfect measure of wages as 5,531 individuals in my sample do not receive compensation of any kind, meaning that they likely resort to subsistence agriculture for their work, and estimating the value of the crops they grow and consume is beyond the current scope of my analysis.

2010/2011 Nepal Living Standards Survey

The second dataset that I use in my analysis is the 2010/2011 round of Nepal's Living Standards Survey[42]. The Nepal Living Standards survey is carried out by the Central Bureau of Statistics and aims to collect data on the measurement of the living standards of people and determine the level of poverty in the country. This is a cross sectional dataset that covers 7,020

[41] *Nepal Earthquake Case Studies*, journeys.dartmouth.edu/NepalQuake-CaseStudies/subsistence-economies-livestock-and-agriculture/#:~:text=Approximately%2084%20percent%20of%20Nepalese,a%20sufficient%20amount%20of%20land

[42] Living Standards Survey 2010-2011, Third Round. Central Bureau of Statistics, https://microdata.worldbank.org/index.php/catalog/1000 . 21 November 2017.

households across 326 primary sampling units across 75 districts. The survey covers a wide range of topics related to household welfare, including demography, consumption, income, access to facilities, education, health, anthropometry. The section most relevant to my analysis was the migration and remittances section. In order to construct the instrumental variable needed for my analysis, I took the average of the share of international migrants of each district in the national living standards survey, which gave me the outmigration rates at the district level. I then merged these values with my main dataset.

Table 1: Summary Statistics of Key Variables

Variable	Mean	SD
Remittance	0.27	0.44
Remittance Amount	198686.8	226741.7
Total Compensation	42.4	120.92
Hours Worked	6.33	2.103
Gender	0.443	0.496
Education	4.68	4.54
Age	39.68	15.52
Household Size	5.4128	2.14
Health Issue Reported	0.136	0.343
Area of Household	5515.9	6542
Cost of Land	1934077	4224012
Total Food Spending	1359	886.67
Amount of Household Loan	83864	202032.7
Land Type: Steep	0.30	0.45
Land Type: Flat	0.445	0.497
Land Type: Moderate	0.252	0.4344
Agriculture	0.911	0.2839
Number of Observations	7722	

There are a total of 7,722 individuals in my dataset. Around 27% of individuals in my dataset belong to a household that has received remittances in the past 12 months, and of those who had received remittances, the average yearly remittance amount was around 198,686 rupees (1653 USD). On average, the individuals in my sample worked 6.33 hours per day, and received a total compensation of 42.4 (0.35 USD) rupees per day. In terms of control characteristics, the average worker in my dataset received 4.68 years of education, and belonged to an average household size of 5 individuals. The predominant sector in my sample is agriculture, employing 91% of the individuals in my dataset.

In order to better understand the underlying control characteristics between individuals who belong to remittance households and individuals who do not belong to remittance households, I created a table of control characteristics by remittance status, where column 1 is the mean of the characteristics of workers who belong to remittance households, and column 2 is the mean of characteristics of individuals who do not belong to remittance households.

Table 2: Descriptive Statistics of Control Characteristics by Remittance Status

VARIABLES	(1) Remittance Mean	(2) Non-Remittance Mean	(3) Mean Difference	(4) p-value difference
Gender	0.397	0.460	-0.063***	0.000
	(0.489)	(0.498)		
Education	4.362	4.805	-0.443***	0.000
	(4.455)	(4.570)		
Age	41.103	39.163	1.941***	0.000
	(16.181)	(15.250)		
Health Issue Reported	0.156	0.129	0.027***	0.003
	(0.363)	(0.336)		
Area of Household	5583.036	5491.096	91.939	0.573
	(6248.685)	(6647.542)		
Remittance Amount	200,000	0.000	200000***	0.000
	(2.3e+05)	(0.000)		
Cost of Land	1800000	2000000	-170000*	0.077
	(3.5e+06)	(4.5e+06)		
Owns Dwelling	0.992	0.983	0.009***	0.001
	(0.090)	(0.128)		
Total Compensation	39.108	43.620	-4.512	0.102
	(99.246)	(128.005)		
Total Food Spending	1349.087	1363.863	-14.775	0.508
	(862.277)	(895.573)		
Household Loan Amount	100000	78000	22000***	0.000
	(210000)	(200000)		
Household Size	5.195	5.493	-0.298***	0.000
	(2.200)	(2.118)		
Steep	0.268	0.315	-0.047***	0.000
	(0.443)	(0.465)		
Flat	0.459	0.440	0.019	0.130
	(0.498)	(0.496)		
Moderate	0.273	0.245	0.028	0.015
	(0.445)	(0.430)		
Agriculture	0.946	0.899	0.047***	0.000
	(0.227)	(0.301)		
Number of Observations	2084	5638	7722	

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

From the table, it is evident that individuals who belong to remittance households are more likely to be female, this difference is 6.3% and is statistically significant from 0 at the 99% confidence level. Additionally, individuals who belong to remittance households tend to be older, with a difference of two years. This is consistent with the literature, as migrant workers from Nepal are predominantly young males, and thus it would make sense that the likelihood of being a “left behind” family member would increase your likelihood of being female and being older. Individuals who belonged to remittance households also were more likely to report having experienced a health issue in the past 12 months, to belong to households with larger loan amounts, to live in steep or moderate land area types, and to work in agriculture, and these results are all statistically significant from 0 at the 99% confidence level. The last two observations are interesting, as they may imply that households who live in steep or moderate land types, where agriculture is less productive and households observe periods of food insecurity in the winter months, may turn to migration as a source of risk diversification by offering an alternative source of income during lean agricultural seasons.The fact that remittance households are more likely to engage in the agricultural sector further confirms this hypothesis, as it may imply that that in areas where there few alternatives to working in the agricultural, households seek migration as the best alternative for income. Interestingly, there were no statistically significant differences in the area of the household, cost of land, dwelling ownership, or total food spending between remittance and non-remittance households.

One of the main issues in estimating a causal effect on the impact of remittances on labor supply is the problem of endogeneity. This is because the decision to have a migrant in a household is not random, but self-selected, and thus migration is likely to be correlated with factors that may impact labor supply decisions among the intensive margin. The two main common sources of endogeneity in the migration literature are omitted variable bias and reverse causality[43]. In the case of remittances and labor supply, there may be significant levels of omitted variable bias as households that decide to have a remittance worker may have

[43]Sasin, Marcin J., and David McKenzie. “Migration, Remittances, Poverty, and Human Capital : Conceptual and Empirical Challenges.” *Policy Research Working Papers*, 2007, doi:10.1596/1813-9450-4272.

underlying characteristics such as health, education, social status, etc. that may be fundamentally different from households that decide not to have a remittance worker. Moreover, there may also be an issue of reverse causality as the current labor supply of the household, such as households that live in areas that are exposed to more natural disasters that affect labor supply in the agricultural setting or households which may induce them to send a remittance worker abroad. My table shows that there are multiple control characteristics that may be correlated with an individual's labor supply decision that is also correlated with the remittance status of a household, such as age, gender, and education. This therefore indicates that running an ordinary least squares regression, even controlling for multiple characteristics, would likely yield biased results. Consequently, I turn to a variety of other estimation strategies, including an instrumental variables approach, a propensity score matching method, and newer machine learning techniques of a random forest estimator and a post-double lasso variable selection regression to try to estimate this relationship between remittances and labor supply net of omitted variables bias.

IV. Methodology

Estimation Strategy & Discussion of Variables

My outcome variable of interest is $Hours\ Worked_i$, which denotes the number of hours worked per day by individual workers members in Nepal. This is a continuous variable between 0 to 15 hours. My independent variable of interest, $Remittance\ Household_i$ is an indicator variable that takes the values 0 or 1, indicating whether an individual belongs to a household that received remittances in the past 12 months. The main control characteristics that I include in my models are age, age^2, gender, education, total compensation, land fixed effects, and sector fixed effects. I chose these control characteristics because I believe they are likely to be factors that also affect an individual's labor supply decisions. However, in the propensity score matching methods and the post-double lasso variable regression methods, I also include a variety of other variables in my analysis.

1. Ordinary Least Squares

The first estimation strategy that I use in order to capture the correlation between remittances and labor supply is an ordinary least squares regression approach. As stated previously, due to the large number of observable underlying differences in control characteristics between individuals who belong to a remittance household and individuals who do not, there is a high likelihood that the two types of workers differ in unobservable characteristics as well. Consequently, the OLS estimator is likely to be subject to omitted variable bias. However, finding the OLS estimator is still useful in allowing us to understand the correlation between the two variables.

(1) $Hours\ Worked_i = \beta_0 + \beta_{OLS} Remittance_i + \beta_2 X' + \epsilon_i$

In this setting, β_{OLS} is the average impact of being in a remittance household on an individual's hours worked. X' is a string of control characteristics that are likely to impact an individual's labor supply decision, including age, gender, age^2, education level, total compensation, household size, area of household, job sector fixed effects, and land fixed effects, and ϵ_i is the error term.

2. Instrumental Variables

Due to the strong likelihood of the presence of omitted variables bias in the OLS estimate, I use an instrumental variables strategy in order to obtain a more causal estimate of the impact of migrant remittances on an individual's labor supply. The instrumental variable that I use is the historical share of international migrants per district calculated from the 2010/2011 Nepal Living Standards Survey. This type of network-based instrument has been used previously by Lokshin and Glinskaya (2004) to study the impact of remittances on women's labor force participation rates, and by Phadera (2019) to study the impact of remittances on overall labor supply. Consequently, I employ the same instrumental variable, but in the setting of rural labor supply on the intensive margin.

The first stage regression for an individual worker i in district j of my instrumental variables analysis is as follows:

$$(2)\ Remittance_{i,j} = \beta_0 + \beta_1 Share\ of\ Migrants\ per\ District\ in\ 2010_j + \beta_2 X'_{i,j} + \epsilon_{i,j}$$

The second stage of my regression would be as follows:

$$(3)\ Hours\ Worked_{i,j} = \beta_0 + \beta_{IV}\ \hat{Remittance}_{i,j} + \beta_2 X'_{i,j} + \epsilon_{i,j}$$

Where $\hat{Remittance}_{i,j}$ is the fitted values from the first stage regression, which are plugged into the second stage regression, and β_{IV} is the 2SLS estimator of the impact of remittances on individual labor supply. X' is a string of control characteristics which include age, gender, age^2, education level, total compensation, household size, area of household, sector fixed effects, and land fixed effects.

There are two conditions for a valid instrumental variable. The first condition is the relevance condition, which is that the instrument must be correlated with the treatment variable of interest, and secondly, the exclusion restriction, which is that the instrumental variable cannot affect the outcome variable through any mechanism other than strictly through impacting the treatment variable.

Using the share of international migrants in 2010 as an instrument for belonging to a remittance household is known as a network-based instrument, which has been widely used in the migration literature. Social and network effects are strong predictors of an individual's decision to migrate, especially in the Nepalese context. This is because having a community of people who speak the same language and share a similar culture in the host country is seen as a lowered psychological cost associated with migration. Additionally, having a stock of existing migrants in the host country is also seen to reduce the economic costs of migration as existing migrants often help new migrants find employment and housing in the new country[44]. Moreover, past studies have shown that Nepalese migrants from the same village often follow their friends and families in the same host country and fill a similar niche in the labor market[45]. As mentioned in the literature review, historical migrant network shares at the district level have previously been used as an instrument for migrant remittances in the Nepalese setting by both Phadera (2014) and Lokshin & Glinskaya (2009). Consequently, I use a similar instrumental

[44] Simpson, N. Demographic and economic determinants of migration. IZA World of Labor 2017: 373 doi: 10.15185/izawol.373
[45] Yamanaka, Keiko. "Changing Family Structures of Nepalese Transmigrants in Japan: Split-Households and Dual Wage Earners ." *Global Networks: A Journal of Transnational Affairs* , vol. 5, no. 4, 14 Oct. 2005, https://doi.org/ https://doi.org/10.1111/j.1471-0374.2005.00123.x.

variable in my analysis of the relationship between rural labor supply and remittances. Whether or not the historical share of migrants at the district level satisfies the relevance condition can also be empirically tested for in my analysis.

While the exclusion restriction cannot be tested empirically, I argue that the share of international migrants at the district level in 2010 are unlikely to impact the current labor supply decisions of individuals in Nepal in any channel other than migrant networks. Similar to Phadera's argument, this is because the percentage of international migrants in 2010 at the district level are unlikely to predict or decide an individual's labor supply decisions made half a decade later in 2016. Consequently, I argue that living in a district that had a higher or lower share of migrants in 2010 is as good as random. Since the share of international migrants per district j in 2010 is correlated with belonging to a remittance household for the individual i, this allows for a quasi-random treatment assignment, which allows me to estimate the local average treatment effect of remittances on labor supply along the intensive margin.

However, while I argue that a migrant-networks instrument would satisfy the exclusion restriction, one challenge to my instrument would be the question of persistence, which is that the underlying structures that influenced the decision to send a migrant worker abroad in 2010, may have persisted over time in 2016 and impact the district level labor supply decisions.

It is important to note that this experimental strategy would allow me to compute the local average treatment effect, and not the total average treatment effect. The local average treatment effect does not measure the treatment for the entire population, but only for a subgroup of the population known as the "compliers", which is the population that was affected by the 2010 variation in outmigration rates per district.

3. Propensity Score Matching

In order to account for this potential bias introduced from validity of the exclusion restriction, I employ a propensity score matching method. Propensity score matching methods are typically used when a large number of control characteristics are available in the sample, and a "propensity score" for treatment participation can be calculated based on those characteristics. The idea behind propensity score matching is, an artificial control group can be constructed through the large availability of control characteristics. Consequently, by controlling for the propensity score for treatment participation, assignment into the treatment group can be interpreted as good as random, and thus the point estimate can be interpreted close to causal.

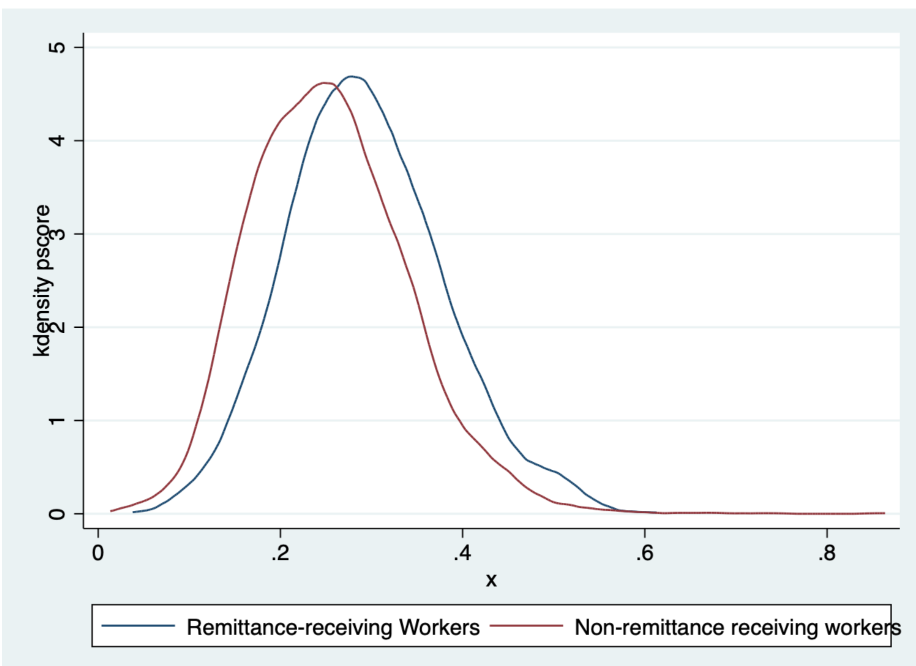
There are two key assumptions of propensity score matching. The first key assumption is that observable control characteristics are believed to be the only factors that affect program participation, and thus datasets with a large number of control characteristics allows for a more accurate calculation of such propensity scores. The second key assumption is that there is a large area of common support, in which treatment observations have a large and roughly equal number of nontreatment-participants with the same propensity score. I argue that propensity score matching methods can be used in my dataset, as the multi-dimensional nature of my dataset allows for a very large number of control characteristics in which an accurate propensity score for belonging to a household that receives remittances.

Under propensity score matching, the probability of treatment conditional on X should be between 0 and 1, and thus a logistic probability model is typically used. Therefore, the first stage of my regression is to compute a propensity score for treatment participation based on a logistic regression:

$$(1)\ Pr(T_i = 1|X_i) = \frac{1}{1+e^{-(\beta_0+\beta_1 X_i)}}$$

In this case, $Pr(T_i = 1|X_i)$ is the conditional probability of belonging to a remittance household, and X_i is a string of control characteristics that may impact belonging to a remittance household, which in my case includes age, gender, education, dwelling ownership, total food spending, total compensation, amount of household loan, age^2, household size, area of household, cost of land, school fixed effects, sector fixed effects, and land fixed effects.

Figure 1: Area of Common Support



From the figure above, it is evident that there is a large area of common support between individuals that receive remittances and individuals that do not receive remittances conditioning on their propensity score. This therefore satisfies the second assumption of propensity score matching.

Consequently, I can estimate the following second stage regression to find the impact of remittances on individual labor supply:

$$(2)\quad Hours\ Worked_i = \beta_0 + \beta_1 Remittance_i + \beta_2 Pr(T_i = 1|X_i) + \epsilon_i$$

Where Hours Worked is the average number of daily hours worked by an individual i, Remittance is an indicator variable of belonging to a remittance household, and $Pr(T_i = 1|X_i)$ is the propensity score for belonging to a remittance household. Therefore, β_1 can be interpreted as the impact of belonging to a remittance household on the number of daily hours worked by an individual, conditioning on the propensity score for remittance participation.

In order to check for robustness, I also estimate the β_1 coefficient with different functional forms of the propensity score. The first is through a cubic functional form of the propensity score such that:

$$Hours\ Worked_i = \beta_0 + \beta_1 Remittance_i + \beta_2 Pr(T_i = 1|X_i) + \beta_3 Pr(T_i = 1|X_i)^2 + \beta_4 Pr(T_i = 1|X_i)^3 + \epsilon_i$$

The second alternative functional form is through inverse probability weighting in which I reweight the treatment group to look like the control group. Consequently, I compute the weights such as that $w(x) = \frac{1}{\hat{P}(x)}$ for the treatment group and $w(x) = \frac{1}{1-\hat{p}(x)}$ for the control group. With these weights for the treatment and control group, I estimate the following regression:

$$Hours\ Worked_i = \beta_0 + \beta_i Remittance_i + \ \delta X_i + \epsilon_i$$

However, one of the limitations of propensity score matching using a logistic propensity score method is that it is a generalized linear model, and thus may be exposed to specification error as it does not capture interactions of covariates or non-linear behavior in predicting treatment assignment. Consequently, I employ a random forest propensity score matching method in order to obtain a better understanding of the relationship between remittances and labor supply.

4. Random Forest Regression Model

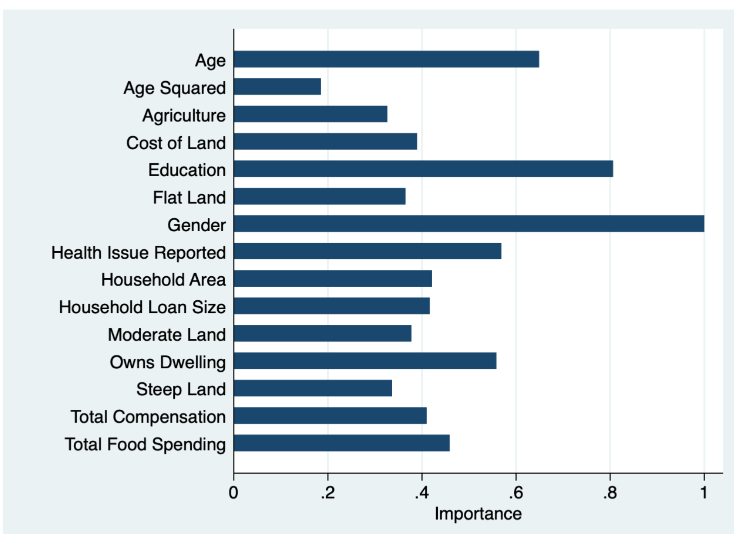
The next regression model in which I use to estimate a causal impact of remittances on labor supply is a random forest estimator as a propensity score. Classification and Regression trees (CART) are built by the recursive partitioning of the data based on different covariates. These decision trees are then bootstrapped and aggregated[46]. The random forests can be used essentially to rank the importance of variables in a regression in a natural way.

The first step in doing so in a dataset is to fit a random forest to the data in which the out of bag error for each datapointed and averaged over the forest. In my analysis, I use around 500 trees

[46] Zhao, Peng, et al. "Propensity Score and Proximity Matching Using Random Forest." *Contemporary Clinical Trials*, vol. 47, 2016, pp. 85–92., https://doi.org/10.1016/j.cct.2015.12.012.

as this is the number of iterations in which the out of bag error becomes stable. The importance score of each covariate is then computed. The propensity score estimated by the random forest regression model therefore allows a better way to compute the propensity of an individual to select into "treatment" given a number of covariates by capturing the best combination of input variables and non-linear effects such as interactions[47].

Figure 2: Rank of Importance of Each Covariate based on Random Forest Model



Based on the figure above, it is evident that based on the random forest regression, the most important covariates that predict an individual's propensity to belong to a remittance household are gender, education, and age. Other covariates that predict remittance include dwelling ownership, and cost of land.

[47] Yahav, Inbal and Shmueli, Galit and Mani, Deepa, A Tree-Based Approach for Addressing Self-Selection in Impact Studies with Big Data (June 5, 2015). Yahav, I., Shmueli, G., & Mani, D. (2016). A tree-based approach for addressing self-selection in impact studies with big data. MIS Quarterly, 40(4), 819-848., Available at SSRN: https://ssrn.com/abstract=2276770 or http://dx.doi.org/10.2139/ssrn.2276770

Therefore, the random forest estimator is able to identify the covariates that best explain the propensity of treatment assignment, in my case, belonging to a remittance household, in a non-linear method. Consequently, by running an OLS regression controlling for this random forest estimator, I am able to obtain the impact of remittances on the number of hours worked. The model is described below:

$$(1)\ Hours\ Worked_i = \beta_0 + \beta_1 Remittance_i + \beta_2 Random\ Forest\ Estimator_i + \epsilon_i$$

Consequently, the variables hours worked and remittance are equivalent to the previous definition, β_1 captures the impact of belonging to a household that receives remittances on the average number of hours worked, and the random forest estimator is the predicted values from the random forest regression model for remittances. I estimate the random forest estimator from the rforest command from the randomForest package in Stata[48].

5. Post Double LASSO Variable Selection Method

The last estimation technique that I use in my analysis is a post-double selection lasso method. This is another machine learning technique, however, instead of using the technique to estimate a propensity score, I use the double selection lasso method in order to understand which covariates to include in my linear regression model. Using a double lasso regression for principled variable selection has been proven to help prevent selecting potentially spurious covariates, and simulations have demonstrated that this method reduces error and increases

[48] Schonlau, Matthias, and Rosie Yuyan Zou. "The Random Forest Algorithm for Statistical Learning." *The Stata Journal: Promoting Communications on Statistics and Stata*, vol. 20, no. 1, 2020, pp. 3–29., https://doi.org/10.1177/1536867x20909688.

statistical power, allowing researchers to statistically control for valid predictors of the dependent variable[49].

The first step in a post double lasso variable selection regression is to fit a lasso regression predicting the dependent variables, and keeping track of the variables with non-zero estimated coefficients. Therefore, in my first stage regression, a lasso regression is run between the hours worked by the individual and a set of control variables.

$$(1)\ Hours\ Worked_i = \beta_1 x_{i.1} + \beta_2 x_{i,2} + ... + \beta_i x_{i,j} + ... + \beta_p x_{i,p} + \epsilon_i .$$

The set of LASSO selected controls can be denoted by A.

In the second step, a lasso regression is run between remittances and the set of control variables.

$$(2)\ Remittance_i = \beta_1 x_{i,1} + \beta_2 x_{i,2} + ... + \beta_j x_{i,j} + ... + \beta_p x_{i,p} + \epsilon$$

The set of LASSO selected controls in this second step regression can be denoted by B.

The third step in a lasso regression is to perform an OLS linear regression of the dependent variable on the independent variable, using the covaries that were selected in the first two steps.

$$(3)\ Hours\ Worked_i = \alpha Remittance_i + w'\beta + \epsilon_i$$

Where $w_i = A \cup B\ i.e.$ the union of selected controls from step 1 and 2

[49] Urminsky, Oleg, et al. "Using Double-Lasso Regression for Principled Variable Selection." *SSRN Electronic Journal*, 2016, https://doi.org/10.2139/ssrn.2733374.

Consequently, through a variable selection through a lasso regression, this allows me to select the relevant control variables through machine learning methods that reduce bias of my point estimate, allowing for a more causal estimate.

‘V. Summary of Results

Table 3: Ordinary Least Squares

	(1) Hours Worked	(2) Hours Worked	(3) Hours Worked	(4) Hours Worked
Remittance	-0.209***	-0.135***	-0.0968**	-0.0956**
	(0.0524)	(0.0520)	(0.0482)	(0.0483)
Gender		0.523***	0.226***	0.230***
		(0.0521)	(0.0486)	(0.0486)
Education		0.00316	-0.0143**	-0.0121**
		(0.00656)	(0.00600)	(0.00602)
Age		0.111***	0.0912***	0.0949***
		(0.00822)	(0.00763)	(0.00767)
Age^2		-0.00133***	-0.00110***	-0.00113***
		(0.0000939)	(0.0000881)	(0.0000884)
Total Compensation		0.000248	0.000606***	0.000483**
		(0.000229)	(0.000190)	(0.000192)
Area of Household		-0.00000909***	-0.00000352	-0.00000447
		(0.00000349)	(0.00000329)	(0.00000330)
Household Size		0.0213*	0.0218**	0.0220**
		(0.0114)	(0.0105)	(0.0104)
_cons	6.392***	4.062***	4.453***	4.116***
	(0.0284)	(0.195)	(1.443)	(1.433)
N	7722	7722	7722	7722
Sector Fixed Effects	No	No	Yes	Yes
Land Fixed Effects	No	No	No	Yes

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The first table summarizes my results for ordinary least squares regression. Column 1 is naive OLS estimating the relationship between remittances and the number of hours worked. The regression coefficient is -0.209 and corresponds to a p value <0.01, and thus is statistically significant from 0 at the 99% confidence level. This can be interpreted as without controlling for other factors, on average, individuals in households that receive remittances on average work 12.54 minutes less per day than households that do not receive remittances. Column 2 controls gender, age, education, age^2 and total compensation, column 3 controls for the previous control variables and includes sector fixed effects, and column 4 controls all the variables and includes both sector fixed effects and land fixed effects. Evidently, as more control variables are included into the regression, it is evident that the magnitude of the coefficient on remittances decreases but remains statistically significant at the 99% confidence level. The regression coefficient in Column 4, with all the relevant control characteristics included, is -0.0956 and corresponds to a p value of 0.05, and thus is statistically significant at the 95% confidence level. This point estimate can be interpreted as, holding gender, age, education level, age^2, total compensation, sector, and land type fixed, an individual belonging to a remittance household works 5.6 minutes less per day than a household that does not receive remittances. Therefore, as more control variables are included in the regression, it is evident that the magnitude of the impact of remittances on labor supply across the intensive margin decreases, but the relationship between the two variables remains statistically significant and negative.

Table 4: Months Worked

	(1) Months Worked	(2) Jul/Aug	(3) Aug/Sep	(4) Sep/Oct	(5) Nov/Dec
Remittance	0.0875	-0.0209*	0.0175	0.0166	0.0410***
	(1.47)	(-2.37)	(1.48)	(1.60)	(3.57)
Gender	0.0733	0.0138	-0.00557	0.00559	-0.00218
	(1.25)	(1.58)	(-0.48)	(0.55)	(-0.19)
Education	0.0209**	0.00231*	0.00777***	-0.00175	0.00505***
	(2.88)	(2.14)	(5.36)	(-1.38)	(3.60)
Age	0.103***	0.00641***	0.00989***	0.00657***	0.0122***
	(11.29)	(4.74)	(5.45)	(4.15)	(6.93)
Age^2	-0.00103***	-0.0000625***	-0.0000919***	-0.0000622***	-0.000133***
	(-9.65)	(-3.96)	(-4.35)	(-3.37)	(-6.49)
_cons	3.922**	-0.256	-0.0642	0.295	0.302
	(2.94)	(-1.30)	(-0.24)	(1.27)	(1.18)
N	7722	7722	7722	7722	7722
Sector Fixed Effects	Yes	Yes	Yes	Yes	Yes
Land Fixed Effects	Yes	Yes	Yes	Yes	Yes

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Due to the seasonal nature of agricultural labor markets, as a robustness check, I also was interested in understanding the difference in labor force participation in different months throughout the year. One of the questions asked in the Household Risk and Vulnerability survey is whether or not the worker worked in different months of the year. My OLS results indicate that controlling for a variety of different variables, on average, there is no difference in the number of months worked in a given year between remittance and non remittance households. However, interestingly, individuals that belong to remittance households are less likely to work in July/August and are more likely to work in November and December. As stated from the figure in Appendix A, July and August corresponds to the lean season in agriculture for Nepal, where food shortages are common. Consequently, this may indicate that remittance households are likely to use migration as a risk diversification strategy, resorting to migration income to combat food shortages in agricultural lean seasons.

Table 5: IV Regression Results

	First Stage (1) Remittance	First Stage (2) Remittance	Second Stage (3) Hours Worked	Second Stage (4) Hours Worked
2010 Share of Migrants	0.288***			
	(0.0298)			
Log 2010 Share of Migrants		0.108***		
		(0.0112)		
Remittance			-3.332***	-2.855***
			(0.550)	(0.523)
Gender	-0.0511***	-0.0523***	0.0663	0.0950
	(0.0117)	(0.0117)	(0.0702)	(0.0666)
Education	-0.00117	-0.00105	-0.0118	-0.0120
	(0.00148)	(0.00147)	(0.00780)	(0.00740)
Age	-0.00635***	-0.00624***	0.0784***	0.0811***
	(0.00179)	(0.00179)	(0.00999)	(0.00948)
Total Compensation	0.0000111	0.0000146	0.000363	0.000370
	(0.0000464)	(0.0000464)	(0.000247)	(0.000234)
Age^2	0.0000924***	0.0000921***	-0.000868***	-0.000910***
	(0.0000208)	(0.0000208)	(0.000121)	(0.000115)
Household Area	0.00000131	0.00000107	-0.00000482	-0.00000495
	(0.000000815)	(0.000000812)	(0.00000430)	(0.00000408)
Household Size	-0.0108***	-0.00976***	0.00112	0.00555
	(0.00247)	(0.00247)	(0.0141)	(0.0134)
_cons	0.0653	0.273	2.774*	2.709*
	(0.310)	(0.310)	(1.649)	(1.564)
N	7153	7153	7153	7153
Sector Fixed Effects	Yes	Yes	Yes	Yes
Land Fixed Effects	Yes	Yes	Yes	Yes

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Due to the possible presence of omitted variables bias and endogeneity concerns that may bias the estimate in the ordinary least squares regression, I run an instrumental variables regression using the share of international migrants in 2010 at the district level as instrumental variable for belonging to a remittance household. Column 1 in table 2 is the first stage regression result. From this regression, it is evident that the coefficient on the 2010 migrant share per district is 0.288, and corresponds to a p value of <0.01, indicating that this estimate is statistically significant from 0 at the 99% confidence level. This therefore indicates my instrumental variable satisfies the first stage relevance condition of instrumental variables, as the share of international migrations at the district level in 2010 is strongly correlated with belonging to a remittance household in 2016. Using 2010 migrant shares per district as my instrument, I estimate the two stage least squares (2SLS) estimator in column 3 to find the local average treatment effect. From column 2, the coefficient on remittance is -3.332. This means that on average, belonging to a remittance household reduces an individual's daily labor supply by -3.332 hours, holding constant variables such as gender, age, education, total compensation, age^2, household area, household size, sector fixed effects, and land fixed effects. This coefficient corresponds to a p value of 0.01 and is statistically significant from 0 at the 99% confidence level.

As a robustness check, I also use the log of the 2010 migrant share by district in Column 2. As stated in the table above, the first stage regression is statistically significant at the 99% confidence level and thus satisfies the relevance condition of instrumental variables. My second stage regression in column 4, corresponds to a result of -2.855, which is slightly less than my regression without using log values of migrant shares, but of similar magnitude of around a 3-hour daily reduction the number of hours worked by individuals who belong to remittance households. Therefore, I conclude that my results are robust to different functional forms of my instrumental variable.

In order to test the relevance of my instrument, I conducted a weak instruments test in Stata[50]. My instrument corresponds to an F statistic of 93.6, and my log of 2010 migrant shares per district instrument corresponds to a F statistic of 93.26. Both of these instruments correspond to an F statistic greater than 10 and thus passes the weak instruments test.

Table 6: Propensity Score Matching Results

	(1) Hours Worked	(2) Hours Worked	(3) Hours Worked
Remittance	-0.131*	-0.127*	-0.130**
	(0.0687)	(0.0683)	(0.0608)
pscore	-2.795***	-22.65***	
	(0.382)	(6.269)	
pscore^2		50.69**	
		(22.16)	
pscore^3		-33.37	
		(24.66)	
_cons	7.119***	9.269***	6.381***
	(0.104)	(0.558)	(0.0430)
N	4684	4684	4684

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

In the table above, I estimate the impact of belonging to a remittance household on the number of hours worked on a typical day, controlling for the propensity score of belonging to a remittance household. Under PSM, the number of observations is reduced from 7,722 to 4,684 because the households that are not in the region of common support are dropped from the

50 See Appendix B

analysis. Column 1 indicates that the coefficient on Remittance is -0.131, and this point estimate is significant at the 90% confidence level. This can be interpreted as, individuals who belong to a remittance household on average reduce the number of hours worked on a typical day by 7.86 minutes. Column 2 estimates this result for a cubic functional form of the propensity score, and column 3 uses an inverse probability weighting of the propensity score. Given these different functional forms of the propensity score, the coefficient on remittance is quite robust, at around -0.13.

As an additional robustness check, I construct balance tables for age, gender, and education at each quintile of the propensity score[51]. In each of the cases, the differences in these underlying covariates are statistically insignificant in 4 out of 5 quintiles of propensity score, indicating that on average, when propensity score is controlled for, the difference between remittance and non-remittance receiving individuals are characteristically very similar.

[51] See Appendix C

Table 7: Estimation based on Machine Learning Techniques

	(1) Hours Worked	(2) Hours Worked
Remittance	-.0190566	-0.0922*
	(.0754214)	(0.0484)
prandomforest	-.2362684	
	(.1299228)	
Gender		0.239***
		(0.0479)
Education		-0.0124**
		(0.00604)
Age		0.0950***
		(0.00742)
Health Issue Reported		0.0745
		(0.0624)
Total Food Spending		-0.0000375
		(0.0000267)
Total Compensation		0.000503***
		(0.000189)
Household Loan		-1.84e-08
		(0.000000110)
Age^2		-0.00114***
		(0.0000862)
Household Area		-0.00000594
		(0.00000363)
Cost of Land		7.14e-09
		(5.84e-09)
Household Size		0.0271**
		(0.0108)
_cons	6.394184 ***	4.071***
	(.040175)	(1.081)
N	4,223	7711

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The table above summarizes the results from machine learning techniques. Column 1 is the regression results from using propensity scores calculated from random forest, and Column 2 is the regression result using the post double lasso variable selection regression. Surprisingly, in column 1, in contrast to the other regression models, when conditioning for the propensity score estimated through a random forest estimator, the coefficient on remittance is -0.019 and is not statistically significant from 0. On the other hand, using the post double variable selection lasso regression method, the coefficient on Remittance is -0.0922, and this estimate is statistically significant at the 99% confidence level. This can be interpreted as, on average, individuals who belong to remittance households on average work 5 minutes less than individuals who do not belong to remittance households.

Table 8: Heterogeneous Gender Impacts

	OLS	IV	PDS Lasso	Random Forest	PSM	PSM Cubic
	(1)	(2)	(3)	(4)	(5)	(6)
	Hours Worked	Hours Worked	Hours Worked	Hours Worked	Hours Worked	Hours Worked
Remit*Male	0.113	3.986***	0.196	.4722819***	0.364***	0.349***
	(0.0980)	(1.209)	(0.123)	(.128)	(0.116)	(0.116)
Remit	-0.147**	-5.114***	-0.249***	-0.209***	-0.317***	-0.300***
	(0.0628)	(0.823)	(0.0843)	(0.117)	(0.0870)	(0.0869)
Male	0.211***	-0.972***	0.197***			
	(0.0545)	(0.325)	(0.0677)			
Education	-0.0136**	-0.00193	-0.0133			
	(0.00605)	(0.00896)	(0.00923)			
Age	0.0941***	0.0906***	0.103***			
	(0.00747)	(0.0113)	(0.00977)			
Age^2	-0.00113***	-0.00103***	-0.00125***			
	(0.0000869)	(0.000140)	(0.000124)			
HH Area	-0.00000545	-0.00000430	-0.00000583			
	(0.00000363)	(0.00000500)	(0.00000416)			
Land Cost	7.02e-09	-8.77e-09	-1.47e-09			
	(5.87e-09)	(8.32e-09)	(6.82e-09)			
Health Issue			0.00398			
			(0.0832)			
Food Spend			-0.0000403			
			(0.0000319)			
Total Comp			0.000550**			
			(0.000251)			
HHLoan			5.67e-08			
			(0.000000130)			
prf				-.2294956		
				(.128)		
pscore					-2.354***	
					(0.398)	
$pscore^2$						-18.06***
						(3.656)
$pscore^3$						29.93***
						(7.432)
_cons	4.249***	3.108*	5.102***	6.393853 ***	7.008***	7.043***
	(1.089)	(1.782)	(1.325)	(.040)	(0.109)	(0.113)
N	7722	7153	7,711	4223	4695	4695
Sector FE	Yes	Yes	Yes	Yes	Yes	Yes
Land FE	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

In order to look at the impact of remittances on women's labor supply, I included an interaction term Remit*Male and reran my models. In the table above, column 1 corresponds to the OLS model, column 2 corresponds to the IV, column 3 corresponds to the post double selection lasso, column 4 corresponds to the random forest regression, column 6 and 7 corresponds to the propensity score matching method with different functional forms of the propensity score. It is evident that the magnitude of the negative coefficient on Remittances increases dramatically throughout the multiple models when compared to the previous results, suggesting that women that belong to remittance households on average reduce their average number of hours worked more than males who belong to remittance households. However, interestingly, in columns 1 and 3, the coefficient on the interaction term is not statistically significant from 0, indicating that there is no difference in the reduction of labor supply between males and females across the intensive margin.

VI. Discussion of Results

Across multiple regression models, it is evident that on average, the relationship between remittances and labor supply across the intensive margin is negative and is statistically significant, with the exception of the random forest propensity score estimator, although of varying degrees of magnitude.

The instrumental variables approach states that on average, individuals who belong to remittance households reduce their daily number of hours supplied to the labor market by 3 hours compared to individuals that do not belong to remittance households. Over the course of one working week, this would imply a reduction of around 15 hours in labor supplied per week.

This is drastically different from the results estimated through the alternative methods, but would be consistent with the current literature in Nepal. This difference in findings may be due to the fact that while my other estimation methods estimate the average treatment effect of remittances on labor supply, the IV model instead estimates the local average treatment effect (LATE). This means that the IV model estimates the impact of remittances on the subset of individuals ("compliers") who are affected by variation of the historical migrant share in 2010 at the district level, rather than the average treatment effect, or the average effect of all individuals who belong to a remittance household[52]. Based on Phadera's finding that belonging to a household that receives remittances on average reduces an individual's labor supply by 8 hours, this would imply that on average, in the agricultural setting, rural households reduce their labor supply by about twice as much as the average individual worker in Nepal. These findings are consistent with the explanation posed by Sharma (2020) that low productivity and wages in the agricultural sector causes farm work to be less preferred, causing individual workers who receive remittances in rural areas to decrease their labor supply to a greater magnitude. Although my findings from the instrumental variables approach are logically consistent with findings estimated by Phadera (2019), the results I have found with alternative models may imply that the instrumental variable of migrant networks that have been used in the Nepalese setting by Phadera (2019) and Lokshin & Glinskaya (2004) may overestimate the impact of migrant remittances on local labor supply. This is because when compared with the average treatment effect, the magnitude of the local average treatment effect estimated by IV is vastly greater in magnitude.

[52] Angrist, Joshua, and Guido Imbens. "Identification and Estimation of Local Average Treatment Effects." 1995, https://doi.org/10.3386/t0118.

On the other hand, across the OLS regression results, the post double lasso variable selection regression, and the propensity score matching methods, the average treatment effect estimate is much lower than the local average treatment effect estimated by IV, as this result to be around 0.10. This implies that on average, individuals that belong to a remittance household reduce their average daily amount of working time by 6 minutes, which over the course of a workweek, would be around 42 minutes. While this may still be economically significant when thinking about the aggregate impacts in this reduction in labor supply across multiple members of a household and across a longer time frame, this effect is significantly less than that estimated by IV. This may imply that in contrast with the previous literature, the average effect of migrant remittances on local labor supply across the intensive margin may not be as drastic as previously thought, and that the income effect and substitution effect resulting from remittances on local labor supply may be more similar in magnitude, though the income effect still dominates greater than the substitution effect. On the other hand, when controlling for the propensity score estimated by the random forest regression model, I find no statistically significant effect of migrant remittances on rural labor supply. Consequently, this would imply that using the random forest propensity score, the income effect is exactly equal to the substitution effect, and thus there would be minimal, or no negative effects on remittances on local labor supply.

My findings also suggest that remittances consistently reduce rural labor supply to a greater extent for women as opposed to men. This suggests that the income effect may be stronger for women than men, as they have a greater propensity for performing household activities than men, resulting in a higher hourly reservation wage, and thus reduce their labor activity more upon receiving remittances. Additionally, women who belong to remittance households may be more likely to be the spouses of migrant workers, and thus are required fulfil more non-work related responsibilities such as taking care of the elderly or their children, and reduce the number of hours supplied to a greater extent than men.

Although I do my best to include robustness checks throughout my analysis through employing multiple functional forms of propensity scores and of the instrumental variable, and conducting my analyses through multiple estimation methods, there are a few limitations in my analysis. Firstly, the lack of wage data due to the large presence of agricultural workers in my dataset may limit the implications of my findings regarding labor supply, as there is no way to estimate the inherent value of the agricultural goods that the workers receive or are compensated by beyond what is reported in the dataset. Additionally, my dataset is also only limited to workers who are currently in the workforce, and thus labor supply can only be estimated along the intensive margin, and not the extensive margin. As mentioned previously, historic migrant shares at the district level as an instrumental variable may be subject to weakness may not satisfy the exclusion restriction if the underlying structures that induced households to send a migrate abroad half a decade ago persist over time, which may explain the large difference in the point estimate in the instrumental variable compared to other estimation methodologies, and thus instruments that can better satisfy the exclusion restriction in the future may allow for a more causal estimate.

Moreover, it may be important to note that while the estimates I find through these analyses are internally valid, they can only be applied to the setting of rural, non-metropolitan areas in Nepal, and cannot be extrapolated to any setting beyond this setting.

VII. Conclusion & Implications for Policy

In order to estimate the relationship between remittances and labor supply, I use a variety of different estimation methods. Using the 2016 Household Risk and Vulnerability Survey, I estimate the impact between remittances and labor supply using an OLS estimation method, an instrumental variables approach using historical migrant shares at the district level, a propensity score matching method, a random forest regression model, and a post double lasso variable selection method. Across different approaches, in general I find a negative and statistically significant relationship between remittances and rural labor supply, with the exception of the random forest propensity score matching method, where I find a negative but statistically insignificant relationship between the two variables. Through the OLS regression, the post double variable selection lasso regression, and the propensity score matching method, I find that, on average, individuals that belong to a remittance household reduce their daily average amount of time supplied to the labor market by 6 minutes. By contrast, the instrumental variables approach shows that remittance-receiving individuals reduce their daily labor supply by 3 hours.

My findings imply that the usage of migrant networks as an instrumental variable, which has been used previously in the Nepalese setting, may overstate the reduction in labor supply in response to remittances. However, the results also consistently suggest that the income effect dominates the substitution effect when it comes to the impact of remittances on local labor supply, as, on average, migrants reduce the number of hours worked in response to receiving remittances. This reduction in labor supply across the intensive margin, while small on the daily level, can eventually be economically significant when aggregated across multiple individuals in a household and across longer time frames. This has implications for Nepal's reliance on foreign countries for its internal economic growth, and implies that remittances may harm the quantity of Nepal's domestic labor market. Nepal's high reliance on remittances leaves it extremely vulnerable to shocks to migration, such as the COVID-19 pandemic which led to a 36.7% loss in remittance income for Nepal in the year 2020[53], as well as the fluctuations of economic conditions in its destination countries, such as oil prices in gulf countries. Consequently, when considering the long term economic growth and development of Nepal, it should consider balancing its welfare gains fueled by remittances, as well as developing its labor force within borders, such as diversifying its sectors by investing in physical and human capital in both agricultural and non-agricultural regions.

Consequently, my findings have several policy implications for Nepal. Given that the income effect dominates the substitution effect in the agricultural sector in response to remittances, to mitigate this reduction in labor supply may require policies to increase the substitution effect. This could result in policies such as investing in improving the productivity of the agricultural sector through mechanization or land irrigation, or investing in more productive sectors. Additionally, policies to reduce credit constraints and encourage self-employment through entrepreneurship in rural areas may be beneficial. Moreover, since remittance households often use migration as a risk diversification strategy, investing in the infrastructure to decrease food insecurity or protecting against environmental vulnerability may also allow households to be less reliant on remittances for their livelihoods, and increase their domestic labor supply.

Avenues for further research may include examining how the amount of remittances received may influence local labor supply in rural areas across the intensive margin, or how

[53]The Impact of COVID-19 on Households in Nepal. Ministry of Agriculture and Livestock Development, Government of Nepal. 2020 September.

migrant remittances impact the sectoral choice of remaining household members. Additionally, the machine learning techniques employed in this paper, such as the random forest propensity score estimator and the post double lasso variable selection regression method, can also be similarly used to analyze the impact of migrant remittances on other areas in development, such as educational attainment of remaining household members or household expenditure.

As migrant remittances continue to fuel Nepal's economic growth and transform its society, it is important to consider both the short term and long term implications on Nepal's economic activity within borders as well. This paper ultimately uses a wide variety of traditional and newer experimental techniques to contribute an insight into the role of remittances in Nepal's story of economic development.

Works Cited

Angrist, Joshua, and Guido Imbens. “Identification and Estimation of Local Average Treatment Effects.” 1995, https://doi.org/10.3386/t0118.

Belloni, Alexandre, Chernozhukov, Victor, Hansen, Christian, Inference on Treatment Effects after Selection among High-Dimensional Controls, *The Review of Economic Studies*, Volume 81, Issue 2, April 2014, Pages 608–650, https://doi.org/10.1093/restud/rdt044

Cannas, Massimo, and Bruno Arpino. “A Comparison of Machine Learning Algorithms and Covariate Balance Measures for Propensity Score Matching and Weighting.” *Biometrical Journal*, 2019, https://doi.org/10.1002/bimj.201800132.

Chami, Ralph, et al. “Are Remittances Good for Labor Markets in Lics, Mics and Fragile States?” *SSRN Electronic Journal*, 2018, https://doi.org/10.2139/ssrn.3221121.

Goller, Daniel; Lechner, Michael; Moczall, Andreas; Wolff, Joachim (2019) : Does the Estimation of the Propensity Score by Machine Learning Improve Matching Estimation? The Case of Germany's Programmes for Long Term Unemployed, IZA Discussion Papers, No. 12526, Institute of Labor Economics (IZA), Bonn

Ghimire, Dirgha, and Prem Bhandari. “Study of Migration and Later Life Health in Nepal.” *Journal of Migration and Health*, vol. 1-2, 2020, p. 100018., https://doi.org/10.1016/j.jmh.2020.100018.

Kharel, Dipesh, From *Lahures* to Global Cooks: Network Migration from the Western Hills of Nepal to Japan , *Social Science Japan Journal*, Volume 19, Issue 2, Summer 2016, Pages 173–192, https://doi.org/10.1093/ssjj/jyw033

Living Standards Survey 2010-2011, Third Round. Central Bureau of Statistics, https://microdata.worldbank.org/index.php/catalog/1000 . 21 November 2017.

Lokshin, Michael, and Elena Glinskaya. "The Effect Of Male Migration For Work On Employment Patterns Of Females In Nepal." *Policy Research Working Papers*, 2008, doi:10.1596/1813-9450-4757.

Murakami, Enerelt, et al. "The Impact of Migration and Remittances on Labor Supply in Tajikistan." *Journal of Asian Economics*, vol. 73, 2021, p. 101268., https://doi.org/10.1016/j.asieco.2020.101268.

Mobarak, Ahmed Mushfiq, and Maira Emy Reimão. "Seasonal Poverty and Seasonal Migration in Asia." *Asian Development Review*, vol. 37, no. 1, 2020, pp. 1–42., https://doi.org/10.1162/adev_a_00139.

Nepal Labor Market Update . International Labor Organization, Jan. 2017.

Nepal Earthquake Case Studies, journeys.dartmouth.edu/NepalQuake-CaseStudies/subsistence-economies-livestock-and-agriculture/#:~:text=Approximately%2084%20percent%20of%20Nepalese,a%20sufficient%20amount%20of%20land

Nepal at a Glance . Food and Agriculture Organization of the United Nations, 2021, https://www.fao.org/nepal/fao-in-nepal/nepal-at-a-glance/en/ .

"Personal Remittances, Received (% of GDP)." *The World Bank* , 2021, https://data.worldbank.org/indicator/BX.TRF.PWKR.DT.GD.ZS?most_recent_value_desc=true.

Phadera, Lokendra. "Impact of International Migration on Labor Supply in Nepal." 2019, https://doi.org/10.1596/1813-9450-9014.

"Remittances." *Migration Data Portal*, migrationdataportal.org/themes/remittances.

Rohwerder, Brigitte. “Seasonal Vulnerability and Risk Calendar in Nepal.” *GSDRC Applied Knowledge Services*, 19 Apr. 2016, https://gsdrc.org/wp-content/uploads/2016/04/HDQ1358.pdf.

Sasin, Marcin J., and David McKenzie. “Migration, Remittances, Poverty, and Human Capital : Conceptual and Empirical Challenges.” *Policy Research Working Papers*, 2007, doi:10.1596/1813-9450-4272.

Siddiqui, Rizwana. “Impact Evaluation of Remittances for Pakistan: Propensity Score Matching Approach.” *The Pakistan Development Review*, vol. 52, no. 1, 2013, pp. 17–44., https://doi.org/10.30541/v52i1pp.17-44.

Sharma, Hari, 2020. "The effect of emigration and remittances on labour supply of the left-behind: Evidence from Nepal," MPRA Paper 102091, University Library of Munich, Germany.

Slavchevska, Vanya, et al. “Rural Outmigration and the Gendered Patterns of Agricultural Labor in Nepal.” 2020, doi:10.2499/p15738coll2.134190.

Schonlau, Matthias, and Rosie Yuyan Zou. “The Random Forest Algorithm for Statistical Learning.” *The Stata Journal: Promoting Communications on Statistics and Stata*, vol. 20, no. 1, 2020, pp. 3–29., https://doi.org/10.1177/1536867x20909688.

Simpson, N. Demographic and economic determinants of migration. IZA World of Labor 2017: 373 doi: 10.15185/izawol.373

Thapa, Sridhar, and Sanjaya Acharya. "Remittances and Household Expenditure in Nepal: Evidence from Cross-Section Data." *Economies* 5.2 (2017): 16. *Crossref*. Web.

The Impact of COVID-19 on Households in Nepal. Ministry of Agriculture and Livestock Development, Government of Nepal. 2020 September.

Urminsky, Oleg, et al. "Using Double-Lasso Regression for Principled Variable Selection." *SSRN Electronic Journal*, 2016, https://doi.org/10.2139/ssrn.2733374.

Walker, Thomas and Jacoby, Hanan. *Nepal Household Risk and Vulnerability Survey 2016*. The World Bank, https://microdata.worldbank.org/index.php/catalog/3705/study-description#metadata-data_collection. 17 September 2021.

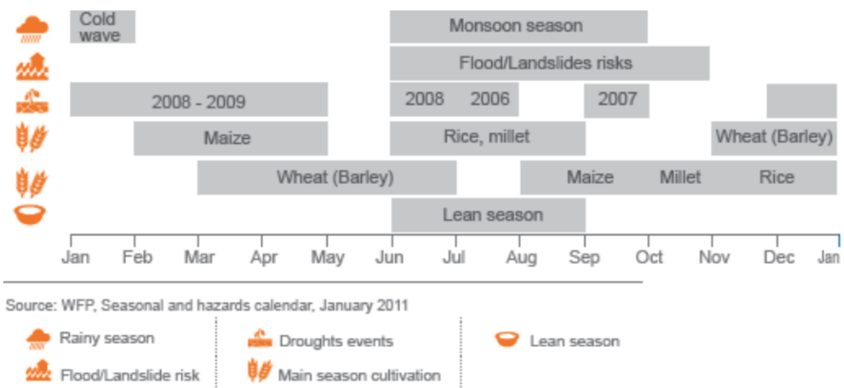
Yahav, Inbal and Shmueli, Galit and Mani, Deepa, A Tree-Based Approach for Addressing Self-Selection in Impact Studies with Big Data (June 5, 2015). Yahav, I., Shmueli, G., & Mani, D. (2016). A tree-based approach for addressing self-selection in impact studies with big data. MIS Quarterly, 40(4), 819-848., Available at SSRN: https://ssrn.com/abstract=2276770 or http://dx.doi.org/10.2139/ssrn.2276770

Yamanaka, Keiko. “Changing Family Structures of Nepalese Transmigrants in Japan: Splot-Households and Dual Wage Earners .” *Global Networks: A Journal of Transnational Affairs* , vol. 5, no. 4, 14 Oct. 2005, https://doi.org/ https://doi.org/10.1111/j.1471-0374.2005.00123.x.

Appendices

Appendix A: Timeline of Critical Seasonal Events[54]

Figure 1: timeline of critical seasonal events



Source: OCHA, 2015, p. 2.

[1] http://drrportal.gov.np/risk-profile-of-nepal Accessed 18.4.16

[54] Rohwerder, Brigitte. "Seasonal Vulnerability and Risk Calendar in Nepal." *GSDRC Applied Knowledge Services*, 19 Apr. 2016, https://gsdrc.org/wp-content/uploads/2016/04/HDQ1358.pdf.

Appendix B: Stata Output of Weak Instruments Test

First-stage regression summary statistics

Variable	R-sq.	Adjusted R-sq.	Partial R-sq.	F(1,7028)	Prob > F
remittance	**0.0508**	**0.0340**	**0.0132**	**93.6857**	**0.0000**

Minimum eigenvalue statistic = **93.6857**

Critical Values # of endogenous regressors: **1**
Ho: Instruments are weak # of excluded instruments: **1**

	5%	10%	20%	30%
2SLS relative bias	(not available)			
	10%	**15%**	**20%**	**25%**
2SLS Size of nominal 5% Wald test	**16.38**	**8.96**	**6.66**	**5.53**
LIML Size of nominal 5% Wald test	**16.38**	**8.96**	**6.66**	**5.53**

Appendix C: Balance Tables Per Quintile of Pscore:

	(1) Age	(2) Age	(3) Age	(4) Age	(5) Age
remittance	-1.876*	1.193	2.048*	-0.787	0.159
	(1.005)	(1.030)	(1.075)	(1.024)	(1.195)
_cons	32.91***	32.77***	32.95***	33.75***	37.45***
	(0.373)	(0.478)	(0.548)	(0.572)	(0.755)
N	937	937	937	937	936

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

	(1) Education	(2) Education	(3) Education	(4) Education	(5) Education
Remittance	0.249	-0.493*	0.0532	0.179	-0.0228
	(0.311)	(0.272)	(0.257)	(0.240)	(0.220)
_cons	8.631***	7.787***	7.168***	7.205***	6.457***
	(0.115)	(0.126)	(0.131)	(0.134)	(0.139)
N	937	937	937	937	936

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

	(1) Gender	(2) Gender	(3) Gender	(4) Gender	(5) Gender
Remittance	-0.0238	-0.0548	0.115***	0.00816	-0.0236
	(0.0289)	(0.0385)	(0.0368)	(0.0331)	(0.0296)
_cons	0.900***	0.634***	0.405***	0.324***	0.278***
	(0.0107)	(0.0179)	(0.0188)	(0.0185)	(0.0187)
N	937	937	937	937	936

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$