LUND UNIVERSITY
School of Economics and Management

Master's Programme in Economic Growth, Populations and Development

Remittances and Household Expenditure in Jamaica

Comparing household expenditure on health and education in remittance receiving versus non-remittance receiving households using Propensity Score Matching

Frankseco Junior Yorke
fyzert@hotmail.com

Abstract: More than 50 percent of Jamaican households are remittances receiving. Remittances generally form a source of supplemental household income by helping households cover expenses such as utilities, healthcare, education etc. Studies have focused on its reaction to economic shocks, structural changes, its use as insurance or safety net or generally just observing trends. This study employs propensity score matching to assess whether there is a significant differences in the expenditure between remittance receiving household (RRH) and non-remittance receiving households (NRRH) in Jamaica using data from the 2015 National Household Survey of Living Conditions. The results show no statistically significant difference in expenditure between the two types of household in the sample assessed. This held true for both outcomes assessed. Households spend more on education than on health irrespective of the household type. RRH households with female heads and household heads with years of schooling in excess of 12 years showed significant positive differences in their expenditure on education. RRH in urban areas were found to be more likely to have higher educational expenditure relative to NRRH. NRRH households were more likely to spend more than RRH on health when at least one member had a chronic illness.

Key words: Remittances, household expenditure, propensity score matching, migration, health, education

EKHS01
Master's Thesis (15 credits ECTS)
June 2019
Supervisor: Kirk Scott
Examiner: Luciana Quaranta
Word Count: 13249

Acknowledgements

The successful completion of this paper would not have been possible without the guidance of my supervisor Kirk Scott who gave valuable advice and guidance throughout the process. I would also like to express gratitude to my family and friends who supported me during this process. Special acknowledgement to my sister Doraine Yorke-Stewart and friends Dwayne, Dane, Sammi, Bryan, Adriana, Jan, and Åsa who stuck with me and provided unwavering support during the writing process, the Swedish Institute that provided funding for this program and finally to God for having sustained me throughout this journey.

Table of Contents

List of Tables

List of Figures

1 Introduction

Remittance receipts has long been a relatively pervasive phenomenon in Jamaica as more than 50% of Jamaican households reported receiving remittances in 2015 (Statistical Institute of Jamaica, 2016). Remittance inflows to Jamaica has seen continued increases recoding over US$2300 million in 2018 (Bank of Jamaica, 2018b). Coupled with tourism it is also one of the largest sources of foreign exchange in Jamaica's service dependent economy (World CIA Factbook, 2018). Researchers have found interest in these remittances coupled with the main factor that has been found to drive remittances – migration. International migration and remittances have been found to significantly reduce poverty levels in the developing world (Adams & Page, 2005).

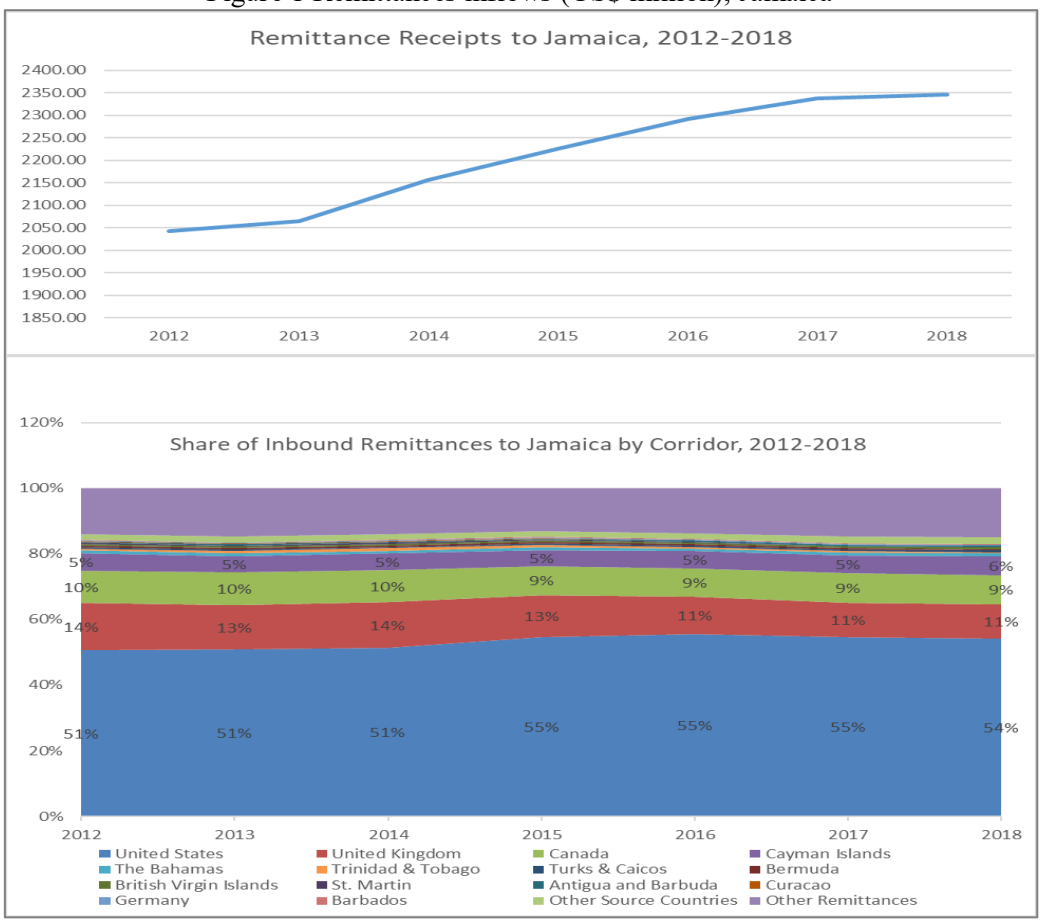
The primary driver of this migration is the perception of improved economic opportunities in the destination countries, strength of familial and social networks in the host country as well as proximity of sending and receiving countries (Thomas-Hope, 2002). As a result migration from Jamaica has historically mainly been to developed countries such as United States, Canada and the United Kingdom. In the case of first two, proximity plays and role, but in the case of the United Kingdom strength of familial and social networks may have a larger impact as Jamaica was once a British overseas territory until late 1900s.

More than 50 per cent of the remittances to Jamaica come from the United States (Henry, Moulton & Ricketts, 2009; Ramocan, 2011). Between 2012 and 2018 (inclusive), the USA remained the largest source market averaging in excess of 50% share, followed by UK (12%), Canada (9%), and Cayman Islands (5%) (Bank of Jamaica, 2018a, 2018b) (See Appendix 1). These remittances generally form a source of supplemental household income helping them to cover expenses such as utilities, healthcare, education etc. The households receiving remittances spend on necessities, mainly education and health with less focus on investments or income-generating opportunities (Asian Development Bank & World Bank, 2018). The share of these remittances among rural and urban areas was almost even with rural areas having a slightly higher proportion (53%) relative to urban areas (47%) (Statistical Institute of Jamaica, 2016).

Jamaica has featured high in both regional and global rankings of remittance receipts. Remittances receipts in Latin America and the Caribbean (LAC) grew in 2017 by 8.7%,

reflecting an unprecedented high of US$ 80 billion (World Bank, 2018b). In 2018 Jamaica ranked 54 in the world's remittance receiving countries receiving over US$2500 million while ranking in the top 10 among its regional counterparts of the LAC (World Bank, 2018a). When a per capita relative is made data shows Jamaica in the top 20 ranked remittances receiving countries globally and holds the #1 rank in the LAC region realizing 17% share of GDP (World Bank, 2018a).

Figure 1 Remittances inflows (US$ million), Jamaica



Source: Bank of Jamaica

The relative magnitude of remittance receipts has expectedly resulted in it being an area of interest for research. Studies on Remittances in Jamaica have focused its contribution to economic development. Increased remittances suggests greater consumption by households which may improve their members' productivity through investment in human capital (health and education) thus leading to development (Samuel, 2004). Studies have examined remittances' reaction to economic shocks (Ricketts, 2011), structural changes (McLean, 2008), its use as insurance or safety net (Clarke & Wallsten, 2003) or generally just observing trends in the remittance receipts (Ramocan, 2011) or of the households that actually received these

remittances. Not much work has been done assessing the comparative expenditure outcomes for households that are not recipient of these receipts. Given the fact that a significant share of Jamaican households receive remittances it may be interesting to assess expenditure in these households and make an attempt at determine whether or to what extent being in such a household may be likely to influences their expenditure. By comparing outcomes from non-remittance receiving households with similar characteristics as those receiving remittances, one may be able to gain addition insight on any likely impact remittances may have on households.

The aim of this study is to assess whether there are differences in expenditure outcomes of households that receive remittances in Jamaica by focusing on health and education spends. To make this assessment we aim to provide a response to the following questions:

1. *Is there a significant difference in the household expenditure on education for remittance receiving households (RRH) versus non-remittance receiving households (NRRH) in Jamaica?*
2. *Is there a significant difference in the household expenditure on health for remittance receiving households (RRH) versus non-remittance receiving households (NRRH in Jamaica?*

In attempting to answer these questions we will explore four major themes; the likely impact of the composition of the household, the characteristics of the head of the households, the geographical location and the likely impact of other socio-economic characteristics on the expenditure on health and education in both types of households. Based on the results of most previous work in this area it is expected that RRH will have significantly greater expenditure based on these characteristics in both cases (health and education). It is also expected that that households, whether remittance receiving or not have significantly greater expenditure on education than on health.

The study employs the propensity score matching (PSM) technique which is a statistical matching technique that makes similar treated and untreated group by accounting for confounding factors observed. The propensity scores will be used to match remittance receiving households with non-remittance receiving households based on confounders for the aforementioned major themes. The assessment of the effect of remittance receipts on household expenditure will be shown through the mean difference in outcomes (health and education expenditure) across the two groups (remittance receiving versus non-remittance receiving

households). The main purpose of this study is to add to the existing literature on how international remittances impact household expenditure patterns with special focus on health and education expenditure while also examining whether the results using this methodology are similar to those used in other studies.

This study is divided in six (6) main chapters. Chapter 1, the Introduction provides a brief background and purpose of the study. Chapter 2 will review theories of migration and remittances and delve into the results of previous related research. The third chapter covers the research design and methodology where challenges and strategies employed will be explored. Chapter 4 will present the data used in this study by providing a descriptive analysis of the data. This will be followed by the presentation and interpretation of the results in Chapter 5. The study will then conclude in Chapter 6 with finalizing remarks and recommendations.

2 Theoretical Framework

2.1 Related Theories

As a result of the inextricable linkage between remittance and migration, many studies have highlighted theories of migration and sought to provide the connecting link to remittances. This study we also will highlight theories of remittances. The literature commonly points to three main theories of remittances. Remittance theories highlight the main motives that migrants have for sending remittances to their sending countries. Understanding these theories offer the benefit of contextually rationalizing the household expenditure patterns of Jamaicans. Lucas and Stark, 1985 posited *three* main motivations to remit. These are self-interest, pure altruism and tempered altruism or enlightened self-interest. These are further discussed briefly.

2.1.1 Remittance Theory: Pure Self Interest

This theory suggests that migrants' decision to remit including the frequency and magnitude of remittance is based on purely selfish motives (Stark & Lucas, 1985). They have expectations of current and future benefits from sending the remittances. Stark & Lucas, 1985 asserts that a pure self-interest is likely when migrants have expectations to return to their sending country. As a result they remit money for one of three reasons;

1. Aspirations of inheritance – pure self-interest remitting migrants remit with the hope of having favour with family back home and hopes that their contribution to the household will increase their likelihood and magnitude of future inheritance.
2. Investment in human and/or physical capital – the migrant's trust in family back home results in him/her remitting so that these investments may be made on their behalf. Examples of these investments include physical capital – land, business starts, cattle etc. Human capital – investment in education for children of the migrant.
3. Investment in fixed capital to ensure migrant's prestige (improved standard of living) upon return to the sending country.

2.1.2 Remittance Theory: Pure Altruism

This theory suggest that migrants have genuine care for the household that is left behind. In this model, the migrant gains utility from the utility of the household in the sending country (Stark & Lucas, 1985). The magnitude of remittance sent is dependent on the migrants' wage level and also on the income level of the household left behind (Samuel, 2004; Stark & Lucas, 1985).

2.1.3 Remittance Theory: Tempered Altruism or Enlightened Self-Interest

Lucas & Stark, 1985 developed this theory as they asserted that neither pure altruism nor pure self-interest sufficiently explained the variability in remittances received. Thus, they developed a theory that suggests that remittances is part of an inter-temporal, mutually beneficial contract between the migrant and the household in the sending country. This theory surrounds matters of investment and risk diversification. The household invests in the human capital of the migrant who is then expected to repay this investment by remitting funds upon migration (Rapoport & Docquier, 2005). The risk diversification may arise where the migrant is expected to help the left behind household in times of crises. This assistance may also be reciprocated in the event that the migrant also encounters crises in the host country (Rapoport & Docquier, 2005).

2.2 Previous Research

2.2.1 Studies on Remittances in Jamaica

A number of studies on Remittances have been undertaken in Jamaica. Their focus have been on the reaction of remittances to economic shocks, structural changes, its use as insurance or safety net or generally just observing trends in the remittance receipts or of the households that actually received these remittances. Much of these have been commissioned by the economists at the Central Bank of Jamaica.

(Diether Beuermann, Ruprah & Ricardo Sierra, 2014) using the 2010 round of the Jamaica Survey of Living Conditions sought to assess the relevance and significance of remittances as a safety net mechanism in Jamaica. They found that shocks adversely affect total household expenditures by an average of 19 percent. However this adverse effect was completely offset by remittances for the households that reported receiving remittances. They also concluded that remittances are more significant in the absence of privately funded insurance by the households (Diether Beuermann, Ruprah & Ricardo Sierra, 2014).

Henry, Moulton & Ricketts, 2009 in a study of the motivation for sending remittances to Jamaica found that there was a combination of altruistic and self-interest/investment motive for sending remittances. Their study showed seasonal variations in remittances where during holiday and back to school periods pointing to evidence of altruism. Additionally, remittance receipts were found to increase significantly in times of natural disasters in the sending country. Their finding suggests that remittances when received would be sent for a specific purpose.

Remittance has a distributive impact on households where remittances are usually used to take care of household basic needs (McLean, 2008). This was the result of a study on the trends in the remittance industry in Jamaica. Similar to Henry, Moulton & Ricketts, 2009 this study also observed seasonal patterns in remittance receipts and concluded that this suggests the existence of a direct link between remittance inflows and seasonal consumption levels giving more strength to the assertion of altruism.

Ramacon 2011 found that the remittance receiving households use over 85 per cent of their remittance receipts on utility payments and basic consumption. For those who received

remittances frequently (at least every other month), more than 50 per cent was used on Utilities, Food and Education combined. – 19%, 18% and 14% respectively.

2.2.2 Remittance and Expenditure on Health

Remittances may impact the health expenditure and outcomes of remittance receiving households. This may be the case especially if the receiving household comprises dependent family members such as children and elderly. (Terrelonge, 2014) asserts that an increase in remittances may be inputs into improved healthcare for the child left behind. These remittances may be with or without prompt by the household left behind. Changes in health conditions or the cost for healthcare may result in requests for greater remittance flows to the household left behind. (Terrelonge, 2014). Given the altruist view and assertion that remittances are sent for specific purpose this suggests that a significant share of remittance receipts may be expected to be allocated to health expenses especially in households with children.

A study including remittance receiving households in Mexico reported that international remittances increase healthcare expenditure (Amuedo-Dorantes & Pozo, 2011). This study also highlighted that expenditure on healthcare was three times more sensitive to variations in remittances relative to changes in other sources of household income. The decision and magnitude of the healthcare spend may be influenced by structural factors in countries. Amuedo-Dorantes & Pozo, 2011 found that lower income households spent less of remittances on healthcare because of mandatory health insurance coverage in Mexico. Valero-Gil, 2008 found a positive significant relationship between remittance income and the proportion of Mexican household expenditure on health particularly among households without medical insurance access. Approximately 10% of changes in remittances in Mexican households in 2004 were devoted to health expenditure (Valero-Gil, 2008).

Remittance receipts have been found to impact nutrition and mortality in children. (Antón, 2010) in his study on the impact of remittances on the nutritional status of children under 5 year old in Ecuador concluded that there was a positive and significant effect of remittances on nutritional status of these children. Another study by Terrelonge 2014, using data from 138 countries, including different regions and covering a 15 year period reported that concluded that increased remittances resulted in reduced child and infant mortality. This study also asserted that remittances is a motor for improved living standards in a country.

While most studies show a positive relationship between remittances and health (Alejandra & Manuelita, 2003) found that parental migration has the possibility of weakening incentives for the accumulation of human capital. This results from the lack of motivation for academic pursuits because of the knowledge that income will come from a family or friend from overseas.

2.2.3 Remittance and Expenditure on Education

Studies on the impact of remittances on education human capital has been found to have varying relationships. Remittances increases primary school enrolment and reduces child labour (Alejandra & Manuelita, 2003). Similar positive relationship was found by (Hines, 2014) who looked at the impact of migration and remittances on household human capital investment in Kenya. He found that remittances increased children's probability of acquiring quality education. Acquisition of higher education has also been found to be influenced by remittances (Arif et al., 2018). Griffith and Rothstein, 2009 highlighted the relationship of remittances and poverty as remittances encourage parents of poor families to send their children to school. They also found that there was a reduction in school dropout hazards rates as a result of increased remittance receipts. (McKenzie & Rapoport, 2011) agreed with this relation as they found that remittances reduced poverty and thus families are more likely to send children to school.

Contradicting results were found in a study in Tajikistan where it was found that emigration has a negative effect on the education of children left behind. It reported that school attendance was inversely related to emigration of family member especially among older children and children from less educated households. It was reported that remittance receipts did not offset this negative effect (Dietz, Gatskova & Ivlevs, 2015). Negative findings were also found in the Latin America and the Caribbean region. A study in rural Mexico found that remittances can have negative effects as it encourages dependence on the relative abroad especially in cases where schooling for the child/ren left behind is not mandatory. These negative outcomes may also arise because of the absence of proper guidance which would have been likely had the guardian not migrated (McKenzie & Rapoport, 2011). This negative effect on motivation to work resulting from remittance inflows to households was also confirmed in a study in Kerala where unemployed persons belonging to remittance receiving households had no urgency to seek employment as they enjoyed the financial support of the emigrant members who remitted funds (Zachariah, Mathew & Rajan, 2001). In Colombia remittances were found to reduced labour force participation because the relative price of leisure was increased by it (Mora, 2013).

3 Methods

3.1 Non-experimental Studies & Biasedness

A recurring problem in non-experimental study analysis is that of overcoming the issue of bias when determining causal relations. Unlike experimental designs where there is random assignment to the treatment and control groups, the non-experimental design may result from bias due to self-selection or bias linked to researcher influence in sample selection (whether intentional or not). When we conduct analyses on remittances in households we need to acknowledge its close link to migration. Due to the non-random selection into migration (Borjas 1994; Chort & Senne 2015; Rooth & Saarela 2007) this may result in biasedness when analyzing the expenditure patterns of remittance receiving households compared to non-remittance receiving households. This bias may result from observable patterns that may not be solely linked to the type of household. Instead the patterns observed may actually be linked to other characteristics of households with high propensity to migrate or other characteristics that are not comparable across the two household types. This may result in differences in the magnitude of remittances and also how such remittances are spent. The likelihood of the existence of such differences may present a challenge when assessing the impact of remittance receipts. Therefore randomization reduces selection and accidental bias. It also allows for comparability where confounders are concerned. The advantage of this is that any differences observed between the two groups can be safely attributed to the treatment and not as a result of underlying confounders.

To address the issue of biasedness resulting from the absence of randomization in non-experimental studies three approaches have been offered and elaborated by (Chiuzan, 2018). Firstly, stratification could be employed where variables that may affect the different groups are split into categories. However small sample sizes may limit this approach especially in cases where there are a number of sub-categories arising from increased number of variables and categories in these variables (Garrido et al., 2014). The number of responses in each category may then be too small to do any robust analysis.

Secondly, adjusted analysis may be employed where adjustments are made for various covariates using regression models. However these models come with a lot of assumptions that many times are violated or unaccounted for, such as linearity, questionable distribution of

residuals or unequal distribution of confounders across the treatment groups. Thirdly, Matching Methods, specifically Propensity Score Matching may be used to address the lack of randomization if the foregoing does not sufficiently offer a solution. Matching involves pairing similar treatment and control units in relation to their observable characteristics. PSM can provide an unbiased estimate of the treatment effect when the pre-treatment covariates captures the differences in the groups in a situation where outcomes are not dependent on assignment into the groups conditional on the pretreatment covariates (Dehejia & Wahba, 2012).

3.2 Propensity Score Matching

Propensity Score Matching (Rosenbaum & Rubin, 1983) is an alternative to the normative adjustment analyses such as regression analyses. The main idea behind propensity score matching is to estimate a random experiment by creating matching sets of participants for treatment and control groups based on their propensity scores (the probability of assignment of a unit with specific characteristics to the treatment group versus the controlled group) (Garrido et al., 2014; Heinrich, Maffioli & Vázquez, 2010; Stuart, 2010; Thoemmes & Kim, 2011). A matched set consists of a unit from the treatment and one from the control group with similar propensity scores. This solves the problem of dimensionality by compressing the relevant factors into a single score (Garrido et al., 2014). By doing this issues surrounding selection bias in non-experimental studies may be reduced or eliminated.

In randomized studies, the propensity score is known. However, in a typical non-experimental study, the propensity score is unknown, because the treatments are not assigned by the researcher. As a result the propensity scores are often times estimated from a logistic regression of the treatment on the units' pre-treatment characteristics. However, the treated and controlled groups may not be directly comparable resulting from systematically differences in the pre-treatment characteristics. Propensity score helps in balancing the groups to make them comparable. The treated and untreated subjects with the same propensity scores have identical distributions for pre-treatment characteristics (Rosenbaum & Rubin 1983).

The steps employed in successfully employing the propensity score matching method for this study are briefly highlighted in the following paragraphs.

3.2.1 Size of Sample

Firstly, a sufficiently large dataset is required. Larger datasets afford increased likelihood of obtaining sufficient matching for analysis. A minimum of 100 matches (200 units in total) may be sufficient for analysis (Stuart, 2010).

3.2.2 Defining the Treatment, Controls and Outcome Variables

The treated, control and outcome variables were defined. In this study, these are remittances receiving households, non-remittance receiving households, and expenditure (health and education) respectively (See table below).

Table 1 Treatment, Control and Outcomes for Propensity Score Estimation

Treatment: Selection in Household Type		
Treated Group	**Control Group**	**Outcome**
Remittance Receiving Households (RRH)	Non-Remittance Receiving Households (NRRH)	Expenditure - Education (ExpEdu) - Health (ExpHealth)

3.2.3 Selecting the Covariates

The covariates of interests were then selected. Since the aim of using this method was to make the groups comparable based on various characteristics, this step is very important. The credibility of the analysis from the propensity score is heavily dependent on covariates selected (Thoemmes, 2012). Their selection should be based on sound theoretical arguments and broader covariates rather than ones chosen on convenience should be used (Shadish, Clark & Steiner, 2008). Ideally we would want to include as many covariates as possible to ensure robustness of the study and achieving comparability. However having too many covariates may result in difficulty in matching (Thoemmes & Kim, 2011). Conversely, having too few covariates might result in biased results. As such as many covariates that were thought to be related to the treatment (Type of Household) and the outcome (Expenditure) were initially included and then by process of elimination reduced based on reference to previous literature, measurability (eg. missing values, coding issues etc.), imbalance assessment (See step 6) and also the need to obtain sufficient matches that would allow for proper analysis. The table that follows summarizes the covariates of interest in the study bases on the four main themes.

Table 2 Covariates used in the study

HOUSEHOLD CHARACTERISTICS	HOUSEHOLD HEAD
Household Size - Members only	Male headed household
All children in the household have their biological father present	Female headed household
All children in the household have their biological mother present	Age of household head
Number of children	Married household head
# of males under 18 years	Years of schooling of household head
# of females under 18 years	Household head - no occupation
Number of children under 6 yrs	Household head Legislators, Senior Officials and Managers
Number of children 6-10 yrs	Household head Professionals
Number of children 11-17 yrs	Household head Technicians and Associate Professionals
# of males below 15 years in household	Household head Clerks
# of females below 15 years in household	Household head Service Workers and Shop and Market Sales Workers
# of individuals age 15-64 dummy	Household head Skilled Agricultural and Fishery Workers
# of females in working age group (15-64)	Household head Craft and Related Trades Workers
# of males in working age group (15-64)	Household head Plant and Machine Operators and Assemblers
# of males 65+ years in household	Household head Elementary Occupations
# of females 65+ years in household	
# of persons in dependent age group in	
LOCATION	**OTHER SOCIO-ECONOMIC**
Urban household	Per Capita Population Quintile
Rural household	# of household members in school
Cornwall County	# of household members in pre-primary
Middlesex County	# of household members in primary
Surrey County	# of household members in secondary
	# of household members in technical or
	# of household members in tertiary
	# of household members in other types of
	Share of household members with health
	Number of chronic illnesses per
	general health score of the household

3.2.4 Estimating the Propensity Scores

Using the selected covariates the propensity scores were then determined by using logistic regression. In this step the aim was to predict the probability of the units (household) receiving treatment (being a remittance receiving household). A logit model was estimated using SPSS where the propensity of each household being assigned to the treated group (RRH) was estimated. This propensity is the conditional probability given the pre-treatment characteristics that have been determined. The true propensity score is unknown and the ideal method of estimation is debatable however logistic regression to date has been the most widely used (Rosenbaum & Rubin, 1983; Thoemmes, 2012). In the logistic regression model, the treatment group (Type of household (RRH or NRRH)) forms the outcome variable while the list of covariates form the independent variable. Simply put, the propensity scores are the probabilities of receiving treatment (being a remittance receiving household) given the set of covariates.

$$e_i = P\ (T_i\ = 1|X_i)$$

Where e = Propensity score
T = Treatment Variable (Household Type) Binary variable that determines whether the observation has treatment
T=1 for treated observations, T=0 for control observations
X = pre-treatment characteristics

3.2.5 The Matching Process

The estimated propensity scores were then used to carry out the matching process where propensity scores for the treated (RRH) and untreated (NRRH) were matched. In theory this could be done using various techniques such as nearest neighbor, inverse probability of treatment weighting, stratification, with and without replacement (Rosenbaum & Rubin, 1983; Stuart, 2010). The nearest neighbor matching was employed in this study as it is the most common and relatively simple technique (Stuart, 2010; Thoemmes & Kim, 2011) in which an untreated unit (non-remittance receiving household) is matched with a treated unit (remittance receiving household) based on having similar estimated propensity scores. To improve the quality of the matches one-to-one matching (Stuart, 2010) was employed. In this situation as single control unit (NRRH) is matched with a single treated unit (RRH). This ensures that the difference in propensity scores between the two groups are minimized and as such reduce bias (Dehejia & Wahba, 2012). The precision of the estimate could be increased by using one-to-many matching, but this may result in increased bias (Dehejia & Wahba, 2012).

Employing nearest neighbor one-to-one matching by itself has its challenges since it is possible that nearest neighbor may have large differences in propensity scores indicating low levels of comparability (Chiuzan, 2018; Pan & Bai, 2015; Rosenbaum & Rubin, 1983; Thoemmes, 2012). To guard against this a caliper which is a maximum allowance in the difference in the matches (Dehejia & Wahba, 2012; Stuart, 2010) was imposed in the matching process. Better balance in the covariates may be achieved with the use of a small caliper, however, this comes with the risk of obtaining fewer matches (Thoemmes, 2012). On the other hand, a larger caliper will yield more matches but create more imbalance and greater bias. The caliper was reduced as much as possible ensuring that the number of matches exceeded 100 (Thoemmes, 2012). This is motivated in the equation below.

A caliper of 0.0001 was initially used. This was adjusted repeatedly until the results seemed to achieve balance in the covariates alongside ensuring that sufficient matches were obtained. A final caliper of 0.005 was used. After iterations of this process a total of 509 matches were found to be sufficient. This reduced the initial sample of 1700 households to 1018 households See table below. In arriving at the matches, those observations with scores outside the areas of common support were removed from the pool.

$$|| p_i - p_j || < c$$

Where each treated observation ***i*** is matched with control ***j*** with minimal difference in the propensity score within the range of common support (defined caliper), ***c***. Five households fell outside the area of common support.

Table 3 Sample sizes after application of propensity score matching technique

Sample Sizes

	All		Matched		Unmatched		Discarded	
Subsamples	Control	Treated	Control	Treated	Control	Treated	Control	Treated
(all cases)	812	888	509	509	301	376	2	3

3.2.6 Checking for Imbalances in Covariates

The "balancing property" of propensity score matching suggest that if we control for the propensity score when the two groups are compared we have effectively transform the observational study into a randomized block experiment, where "blocks" are groups of subjects with the similar propensities (Rosenbaum & Rubin, 1983).

Checks to assess the robustness of the matched results were undertaken to ascertain whether balance on the covariates had been achieved through the matching process. This was done by analyzing the differences in standardized mean and the variance ratio. Standardized mean differences of the covariates close to 0 post matching and variance ratios close to 1 is indicative of balance on the covariates (Thoemmes, 2012). It is important that this step be iterative to ensure that the most robust matches possible are obtained (Rosenbaum & Rubin, 1983).

3.2.7 Analysis using the New Matched Dataset

Finally a new dataset was created using the matched households only. This new dataset was then used to estimate the treatment effect by doing an outcome analysis. This analysis may be done through the use of independent T-tests and ANOVA tests. There has been debate surrounding whether matched data methods (paired tests) should be used (Stuart, 2010). This study did not take that approach on the premise that pairs were not selected on outcome. Notwithstanding the argument may be put forward that the selected pairs are similar leaving an argument for the use of pair sample t-tests. After the nearest neighbor matching was done, we were then able to use the matched non-remittance receiving households and assess the average difference in their expenditure pattern relative to the remittance receiving households.

Following are some of the main assumptions of this matching process:

Conditional Independence: The outcomes are independent of treatment in the non-experimental study as opposed to experimental (random) studies where outcomes are dependent on treatment.

Unconfoundedness: The assignment into treatment doesn't affect the control group's outcome

Matching Assumption: There are treated and control observations for every x value. There is a corresponding matched control observation with similar x characteristics for each treated

observation. The treatment and control group cannot be entirely dissimilar. Some level of overlap is required.

Balancing Condition: Given the same propensity score there will also be similar x characteristics and the assignment into treatment is not dependent on these characteristics. This balancing condition should be testable.

3.3 Data Analysis

Three (3) Statistical Software were used in this study. STATA, Statistical Packages for the Social Sciences 24 (SPSS) and R. SPSS was used because it aided in the management of the datasets as a total of 14 dataset had to be merged. The R Add-in for SPSS was used to establish the matches and generate the new dataset for analysis.

4 Data

4.1 Source Material

Secondary data was be gathered from the 2015 round of the Jamaica Survey of Living Conditions (JSLC). This is cross-sectional data from a national household survey conducted annually in Jamaica. It provides baseline measures of household welfare used to monitor the impact of Jamaica's Human Resources Development Program on health, education and nutrition. Modules in this survey covers indicators such as health, education, anthropometric measurements for children, daily expenses, consumption expenditures, non-consumption expenditures, food expenses, consumption of home production, housing, an inventory of durable goods, other household income, food stamps and employment (World Bank, 2002). The choice of 2015 was based on the fact that it is the most recently available dataset that could be used for the purpose of this study. This round of the survey was based on a 0.3 per cent sample of Jamaican households. This translated into 1716 households and 5154 individuals. For the purpose of this study only international remittances were considered. A total of 894 (52%) household reported that they received remittances in the period while 822 (48%) said they were not recipients of international remittances. This sample was further trimmed to 1700 households. The trimmed sample maintained similar ratios in household types (888-RRH, 812-NRRH). The sample was trimmed because of missing values for some variables and/or seeming outliers for some variables. Some of the variables which had missing values were not covariates used in the study. However they still had to be removed as the PSM method doesn't allow for missing data points – even if they are not from one of the variables used in the matching process (Thoemmes, 2012).

Remittances cover more than personal transfers. It includes the net income from migrant short-term workers net of the expenses for travel, transport, taxes, and social contributions incurred abroad (ed. International Monetary Fund, 2014). Three main remittance measures are put forward by the manual, namely, personal remittances, total remittances and total remittances and transfers to Non-Profit Institutions Serving Households. Remittances can also arise from inter and intra country inflows and outflows. For the purpose of this study focus was placed on inflows of inter country personal transfers (international remittance inflows). Therefore the remittances in this context is defined as the private monetary transfers to households from friends or relatives who live outside of Jamaica.

The JSLC data was provided in a somewhat detached form. Results from each section of the JSLC questionnaire were provided in different datasets. Some datasets were household data, some individual/person and in both cases there were instances of each household or individual having multiple itemed responses from the questionnaire (See Appendix A). This warranted some data manipulation, which included calculations, recoding of variables, computation of new variables and merging of datasets (See Appendix G for Syntax of all new variables, recoding and grouping). A total of fourteen (14) datasets had to be merged to create the final dataset. Four of these datasets were household datasets and were taken as is from the data obtained from the Sir Arthur Lewis Institute of Social and Economic Studies (University of the West Indies (Mona) Jamaica. The remaining ten (10) were datasets that were created by either using, recoding or creating variables from person datasets and converting to household data using the aggregate function in SPSS. All datasets provided in the form of person data had to be manipulated to be able to arrive at a final dataset of households with data relevant to this study (See Appendix F).

A single variable capturing remittance receiving versus non-remittance receiving households was not explicitly provided in the dataset. This was arrived at using variables from Part K of the JSLC Questionnaire (See Appendix B). Once the household had a positive response to at least one of the questions below they were classified as remittance receiving. Conversely, if they didn't have a positive response to at least one of the questions then that household was classified as non-remittance receiving.

During the past 12 months, has any member of your household received income in cash or in kind from the following sources?

1. *Support for children from parents who live abroad*
2. *Spouse/ Partner who lives abroad*
3. *Child / children who lives / live abroad*
4. *Other relatives or friends who live abroad*

For the outcome variables (expenditure on health and expenditure on education), per capita calculations were made. Therefore the expenditure (health and education) refers to the expenditure on health or education per household member. In the case of expenditure on education, the denominator was the number of household members who reported being engaged in some level of educational activity during the period of the survey. This approach was taken to support comparability across households.

4.2 Descriptive Statistics

The sample of 1700 households is summarized in the table below. The table gives a comparison of the characteristics observed for the remittance receiving households versus the non-remittance receiving households utilizing a simple t-test for difference in means. The characteristics were grouped into four broad headings relating to; the general household, household head, location and other socio-economic characteristics.

4.2.1 General Household

Remittances receiving households were more likely to be larger than non-remittance receiving households. They tend to have more individuals of working age. These working age individuals were also more likely to be males. Households receiving remittances were also more likely to have more girl children relative to those not receiving remittances. As it relates to the presence of biological parents, remittance receiving households were less likely to have both biological parents present for all children within the household. Note however, that the presence of the biological mother was found to be insignificant. This is in line with anecdotal assertions of the prevalence of absentee fathers in Jamaica. Other observed general household characteristics though insignificant, were; Remittance receiving households were more likely to have children and elderly women (65+ years).

Table 4 Descriptive statistics of selected household variables of remittance receiving and non-remittance receiving households

	Remittance Receiving		Non-Remittance Receiving		Difference
	Mean	**Std. Dev**	**Mean**	**Std. Dev**	
HOUSEHOLD CHARACTERISTICS					
Household Size - Members only	3.074	2.072	2.905	1.894	0.169***
All children in the household have their biological father present	0.105	0.306	0.135	0.342	-0.031***
All children in the household have their biological mother present	0.252	0.435	0.256	0.437	-0.004
Number of children	0.998	1.330	0.959	1.290	0.038
# of males under 18 years	0.519	0.869	0.514	0.843	0.006
# of females under 18 years	1.124	0.897	1.007	0.857	0.116*
Number of children under 6 yrs	0.312	0.620	0.302	0.617	0.01
Number of children 6-10 yrs	0.245	0.542	0.264	0.527	-0.018
Number of children 11-17 yrs	0.440	0.768	0.394	0.707	0.046
# of males below 15 years in household	0.419	0.767	0.435	0.767	-0.016
# of females below 15 years in household	0.382	0.709	0.381	0.704	0.001
# of individuals age 15-64	1.966	1.352	1.808	1.237	0.158*
# of females in working age group (15-64)	0.920	0.850	0.884	0.773	0.036
# of males in working age group (15-64)	1.046	0.986	0.924	0.914	0.123*
# of males 65+ years in household	0.133	0.343	0.133	0.343	0
# of females 65+ years in household	0.175	0.394	0.149	0.367	0.026
# of persons in dependent age group in household	1.108	1.227	1.097	1.177	0.011

* 10 %; ** 5 %; *** 1 %

4.2.2 Household Head

Looking at the characteristics of the household head, remittance receiving households were dominated by female heads relative households not receiving remittances. These households were more likely to have heads who do not have an occupation. This may be an indication that receipts from remittances may act as a form of income and may be used to cover these households' basic expenses. A somewhat surprising observation was that heads of remittance receiving households were less likely to have occupations classified as elementary occupations (the lowest on the scale from the ILO's International Standard Classification of Occupations). However, it could be argued that this could be explained by the fact that these heads are more likely to have more years of schooling thus enabling economic advantage, albeit that observation was insignificant.

Other observations relating to the household heads in RRH, though insignificant are that their household heads were more likely to be older (mean 52 years), married and have more years of schooling.

Table 5 Descriptive statistics of selected household head variables of remittance receiving and non-remittance receiving households

	Remittance Receiving		Non-Remittance Receiving		Difference
	Mean	**Std. Dev**	**Mean**	**Std. Dev**	
HOUSEHOLD HEAD					
Male headed household	0.515	0.500	0.564	0.496	-0.049**
Female headed household	0.485	0.500	0.436	0.496	0.049**
Age of household head	52.130	16.142	51.016	16.067	1.113
Married household head	0.255	0.436	0.235	0.424	0.019
Years of schooling of household head	11.739	3.695	11.674	3.567	0.065
Household head - no occupation	0.209	0.407	0.159	0.366	0.051***
Household head Legislators, Senior Officials and Managers	0.050	0.217	0.067	0.249	-0.017
Household head Professionals	0.060	0.237	0.064	0.245	-0.004
Household head Technicians and Associate Professionals	0.043	0.203	0.041	0.198	0.002
Household head Clerks	0.062	0.241	0.047	0.211	0.015
Household head Service Workers and Shop and Market Sales Workers	0.136	0.343	0.127	0.333	0.009
Household head Skilled Agricultural and Fishery Workers	0.173	0.379	0.188	0.391	-0.015
Household head Craft and Related Trades Workers	0.111	0.315	0.119	0.325	-0.008
Household head Plant and Machine Operators and Assemblers	0.053	0.224	0.042	0.200	0.011
Household head Elementary Occupations	0.102	0.303	0.147	0.354	-0.044***

* 10 %; ** 5 %; *** 1 %

4.2.3 Location of Household

When location of the household is examined, the rural/urban phenomenon did not show significant results. However there seems to be a likelihood for remittance receiving household to reside in urban areas. Significance was however observed for the county in which the household was located. Jamaica has three counties, Cornwall, Middlesex and Surrey (See Appendix C). Remittance receiving households were more likely to be located in the county of Middlesex while non-remittance receiving households were more likely to be located in Surrey.

Table 6 Descriptive statistics of selected location variables of remittance receiving and non-remittance receiving households

	Remittance Receiving		Non-Remittance Receiving		Difference
	Mean	Std. Dev	Mean	Std. Dev	
LOCATION					
Urban household	0.518	0.500	0.495	0.500	0.023
Rural household	0.482	0.500	0.505	0.500	-0.023
Cornwall County	0.235	0.424	0.261	0.439	-0.026
Middlesex County	0.507	0.500	0.442	0.497	0.065***
Surrey County	0.258	0.438	0.297	0.457	-0.039**

* 10 %; ** 5 %; *** 1 %

4.2.4 Other Socio-Economic Characteristics

Observation of the other socio-economic characteristics, relative to non-remittance receiving household, remittance receiving households were more likely to be in higher population quintiles. They were also more likely to have more members undergoing secondary, technical or vocational and tertiary schooling during the reference period of the survey. However, as the number of members in primary education increased, households were less likely to receive remittances.

Observations of the other socio-economic characteristics showed that remittance receiving households were more likely to have higher shares of chronic illnesses, they had higher shares of members with health insurance coverage and also had a better general health scores. General health score was calculated for this study by taking the mean of the self-reported health score (1-5) of the household members (See calculation for variable *genhealth* in Appendix G p70.)

Table 7 Descriptive statistics of selected household variables of remittance receiving and non-remittance receiving households

	Remittance Receiving		**Non-Remittance Receiving**		**Difference**
	Mean	**Std. Dev**	**Mean**	**Std. Dev**	
OTHER SOCIO-ECONOMIC					
Per Capita Population Quintile	3.547	1.359	3.265	1.455	0.283***
# of household members in school	1.100	1.390	1.017	1.295	0.083
# of household members in pre-primary school	0.318	0.632	0.307	0.629	0.011
# of household members in primary school	0.321	0.664	0.377	0.672	-0.056*
# of household members in secondary school	0.315	0.639	0.244	0.528	0.071**
# of household members in technical or vocational school	0.028	0.172	0.011	0.105	0.017**
# of household members in tertiary school	0.090	0.344	0.054	0.247	0.036**
# of household members in other types of school	0.028	0.179	0.025	0.155	0.004
Share of household members with health insurance	0.196	0.353	0.175	0.344	0.02
Number of chronic illnesses per household member	0.231	0.506	0.195	0.459	0.036
general health score of the household	2.011	1.852	1.996	1.812	0.015

* 10 %; ** 5 %; *** 1 %

5 Empirical Analysis

5.1 Results & Discussion

5.1.1 Unadjusted Relationship

A simple OLS regression of to examine the correlation between the type of household and their expenditure of education in one instance and their expenditure on health in another instance (equivalent to a t-test) revealed a positive significance difference. Testing per capita expenditure on education relations, shows that the differences in per capita mean spending between the two household types ($3069.67) is significant at the 5% level. This suggests that remittance receiving households on average spend $3069.67 more on education for each household member than households that do not receive remittances. Assessing the per capita expenditure on health revealed a significant difference in the mean (1798.639) between the two household types at the 5% level. This would suggest that remittance receiving households spend on average $1798.64 more on health for each household member than non-remittance receiving households do.

Figure 2 Regression of Household type (independent) and Expenditure on Education and Health (Dependent variables)

Source	SS	df	MS	Number of obs	=	1,700
				F(1, 1698)	=	4.90
Model	3.9967e+09	1	3.9967e+09	Prob > F	=	0.0271
Residual	1.3862e+12	1,698	816394498	R-squared	=	0.0029
				Adj R-squared	=	0.0023
Total	1.3902e+12	1,699	818266380	Root MSE	=	28573

per_Educat~l	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
HHOLDTYPE2	3069.671	1387.363	2.21	0.027	348.551	5790.792
_cons	7758.89	1002.702	7.74	0.000	5792.227	9725.552

Source	SS	df	MS	Number of obs	=	1,700
				F(1, 1698)	=	4.34
Model	1.3722e+09	1	1.3722e+09	Prob > F	=	0.0374
Residual	5.3704e+11	1,698	316276618	R-squared	=	0.0025
				Adj R-squared	=	0.0020
Total	5.3841e+11	1,699	316898097	Root MSE	=	17784

per_Health	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
HHOLDTYPE2	1798.639	863.5222	2.08	0.037	104.9596	3492.319
_cons	7368.719	624.1019	11.81	0.000	6144.629	8592.808

5.1.2 Results from Matching Exercise

After matching on the covariates listed in the previous section a total of 509 matches were found. This translated into a total of 1018 households which now made up the new sample on which the analysis id done. A total of 5 households were discarded as they fell outside the area of common support.

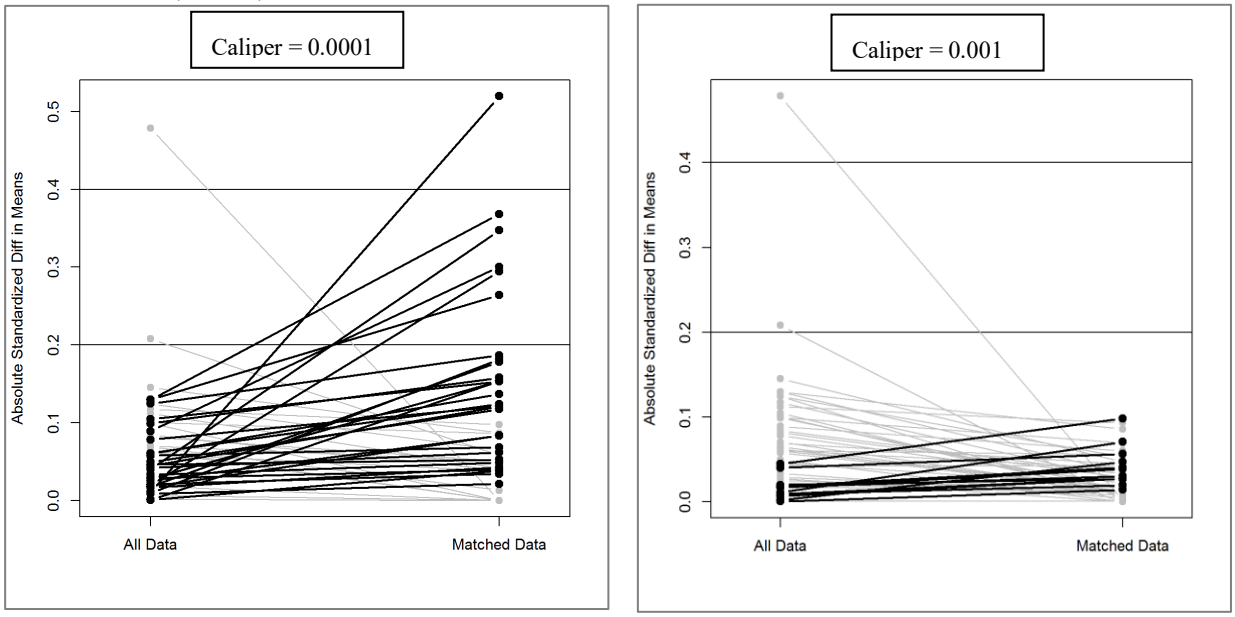
Table 8 Sample sizes of RRH and NRRH post matching exercise

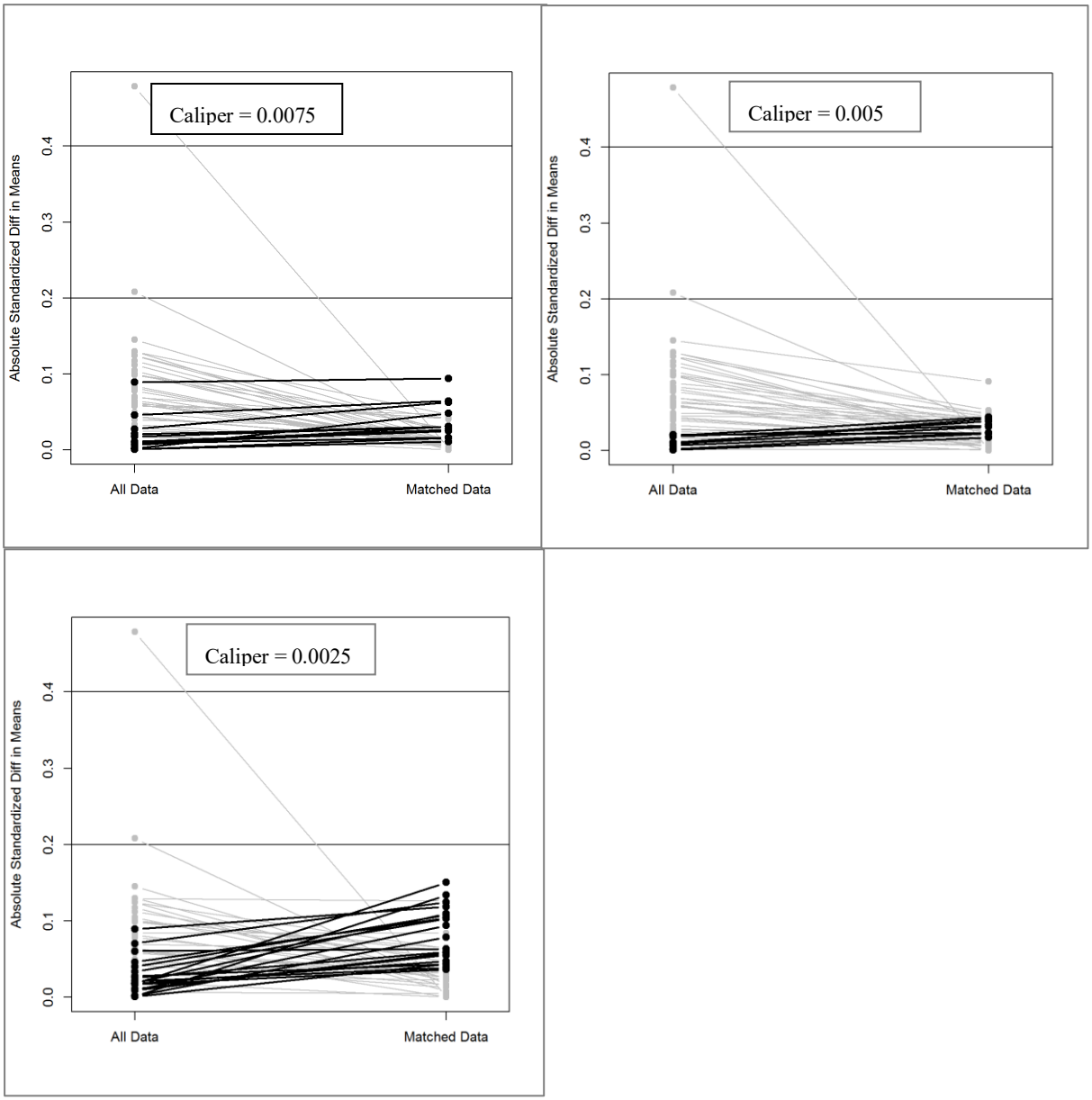
Sample Sizes

	All		Matched		Unmatched		Discarded	
Subsamples	Control	Treated	Control	Treated	Control	Treated	Control	Treated
(all cases)	812	888	509	509	301	376	2	3

The choice of this matched dataset versus others resulted from it having the best results from the balancing checks that were done. The figures below show the line plot of standardized differences in the means before and after matching using different calipers. Note, ideally the best results are achieved when the standardized differences tend towards 0. Dark lines indicate an increase in the differences post matching which is an unfavorable result. A caliper of 0.005 yielded the best results.

Figure 3 Standardized difference in means before and after matching, Iterations using Calipers 0.0001, 0.001, 0.005, 0.0075 and 0.0025




Caliper = 0.0075
Absolute Standardized Diff in Means
0.0
0.1
0.2
0.3
0.4
All Data
Matched Data
Caliper = 0.005
Absolute Standardized Diff in Means
0.0
0.1
0.2
0.3
0.4
All Data
Matched Data
Caliper = 0.0025
Absolute Standardized Diff in Means
0.0
0.1
0.2
0.3
0.4
All Data
Matched Data

Appendix D provides the details of the results for all the five (5) tested calipers. A summary of the imbalance check for the chosen caliper of 0.005 is provide here.

Table 9 Overall balance test (Hansen & Bowers, 2010)

	chisquare	df	p.value
(all cases)	14.885	38.000	1.000

The Hansen & Bowers χ^2 Test for overall imbalance was used to assesses whether variables were significantly unbalanced post matching by examining the covariates used in the estimation of the propensity scores

Table 10 Relative multivariate imbalance $\mathcal{L}_1$ (Iacus, King, & Porro, 2010)

	Before matching	After matching
(all cases)	.999	.998

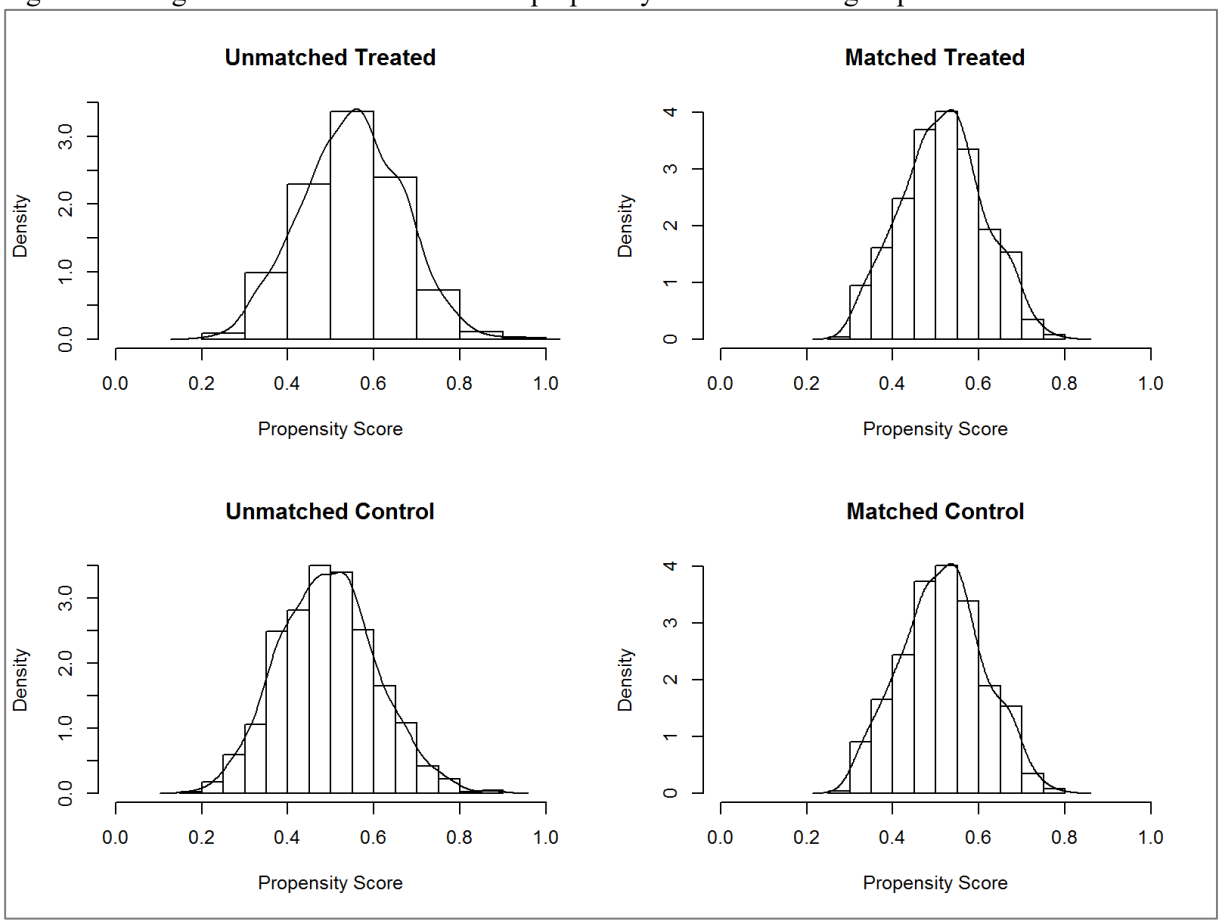
The $\mathcal{L}_1$ measure is based on an assignment of all variables into bins and then doing a comparison of differences in frequencies of all cells of a multivariate contingency table of the two groups (control and treated) (Thoemmes, 2012). The result ranges from 0 (perfect balance) to 1 (complete separation in the cross-tabulation) (Thoemmes, 2012). Theomme, 2012 asserts that $\mathcal{L}_1$ measures are commonly close to 1 when there are many covariates. The key is to ensure that the $\mathcal{L}_1$ post matching is less than $\mathcal{L}_1$ pre- matching (Thoemmes, 2012) which was the case when a caliper of 0.005 was used in this matching, suggesting that matching improved the overall balance.

Table 11 Summary of unbalanced covariates (|d| > .25)

No covariate exhibits a large imbalance (|d| > .25).

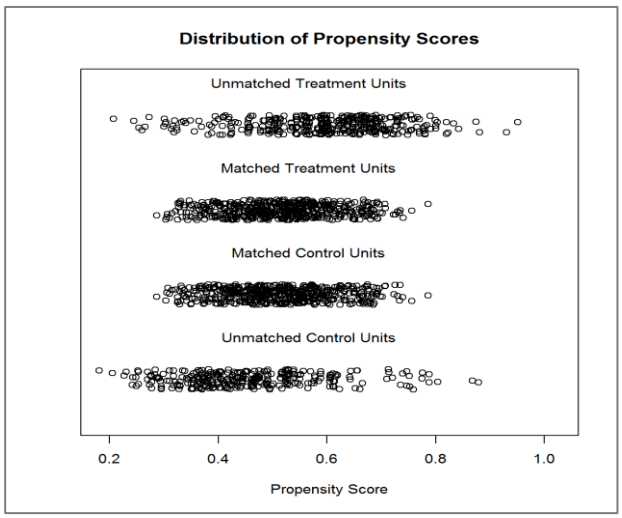
The univariate tests for imbalance in each covariate and every possible interaction shows that there were no as covariates that exhibited a standardized mean difference greater than 0.25 post matching. Therefore there was no need to re-specify propensity scores in order to achieve balance.

Figure 4 Histogram of the distribution of the propensity scores for both groups



A histogram was used to examine the distribution of the propensity scores for both groups. For both the RRH (treated) and the NRRH (control) it appears that matching has made the distribution slightly more normal. Note that the tails of the distribution of the propensity scores were narrowed post matching resulting in a more compressed distribution. Notwithstanding the scores appear to be normally distributed pre and post matching.

Figure 5 Dot-plot of the individual propensity score for both groups



A dot-plot of the individual propensity score provides as similar graphical representation of the distribution pre and post matching as with the histogram. In the dot-plot the presence of outliers can more easily be observed by the clear circles in the unmatched groups. These clear dots become less in the matched groups.

The figure below shows a histogram of the standardized differences of all covariates before and after matching. Recall that a zeroing of the difference is favorable. Post matching it was observed that the standardized differences were notably minimized and were almost centered around zero.

Figure 6 Histogram of the standardized differences of all covariates before and after matching.

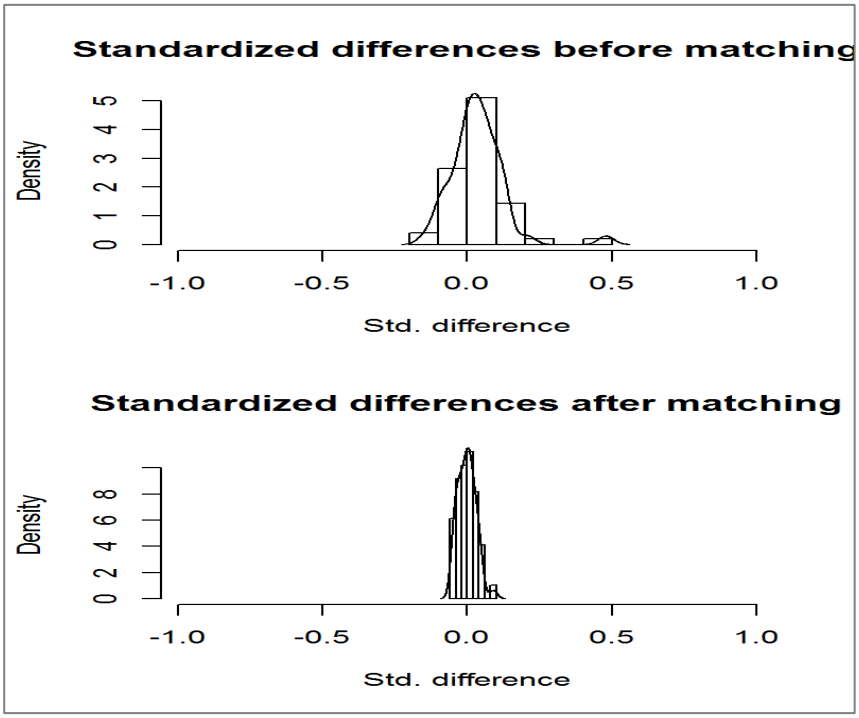
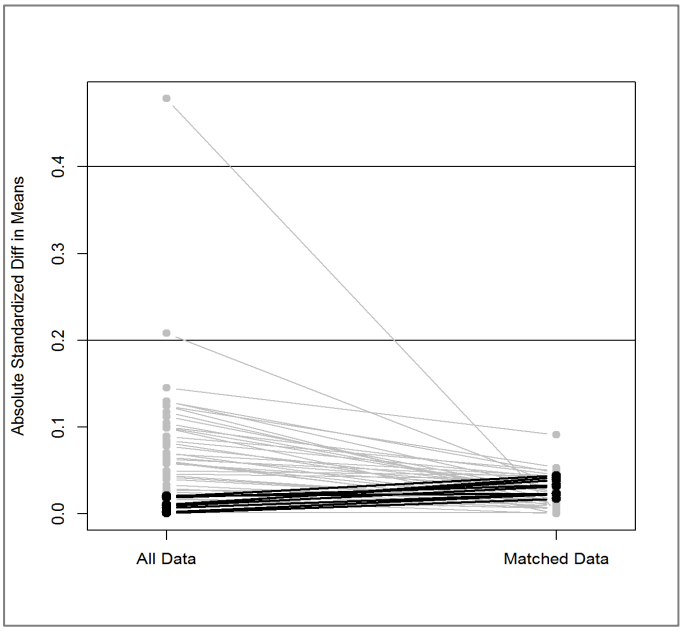
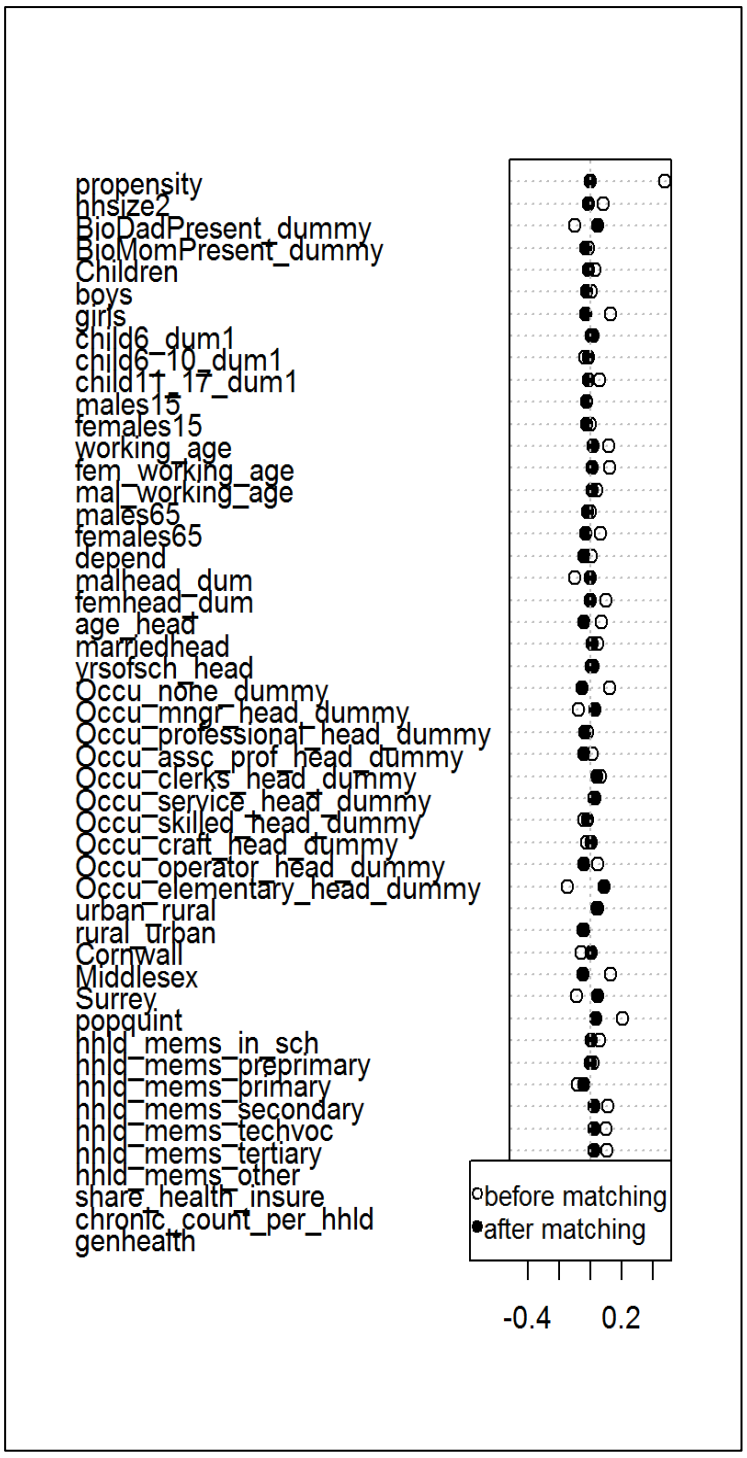


Figure 7 Parallel line plot pre and post matching



The parallel line plot shown illustrates the magnitude of the differences in the standard differences in the propensity score pre and post matching. There were a number of covariates in which there was more imbalance post the matching (indicated by the dark lines). The differences however were relatively small. The dot-plot overleaf provides a similar illustration by listing each covariate used in the matching.

Figure 8 Individual dot plot of covariates



5.1.3 Analysis of New Dataset (Matched Dataset)

Post matching t-tests were used to assess whether there were differences in expenditure based on the type of households. The results are summarized in the table that follows.

Table 12 Analysis of expenditure outcomes post matching

	Remittance Receiving		Non-Remittance Receiving		Difference
	Mean	**Std. Dev**	**Mean**	**Std. Dev**	
Area of Expenditure					
per Capita Expenditure of Education	35 726.66	55 053.79	32 490.40	53 828.60	3 236.26
per Capita Expenditure of Health	7 903.28	16 454.51	8 200.24	17 370.53	-296.96

For the households that were matched, no statistically significant difference was found between the two types of household neither in their expenditure on education nor their expenditure on health. The data showed that on average the difference in expenditure on education between RRH and NRRH was $3236.26 per member more in RRH. The difference in expenditure on health between RRH and NRRH was $296.96 per member less in RRH. Although these general results were found to be insignificant, the relative magnitude of the values may be indicative of the relative importance of education in Jamaican households studied. This observation may be linked to innate desire to invest in educational human capital as well as legislative frameworks that exist relating to mandatory schooling pre the post-secondary level.

The results are generally contrary to many previous studies utilizing other methods. However, the directional relation could be said to agree to some extent with Alejandra & Manuelita, 2003 whose findings suggested a possible weakening of the incentive to accumulate human capital as a result of parental migration. Albeit this studied didn't make a distinction between parental and other sources of remittances.

5.1.4 Matched Data: General Household Characteristics

When the selected household characteristics were examined for the two types of households, no statistically significant difference was observed in the expenditure on health nor on education. Though not significant, the matched data suggests that RRH were in most cases (13 of 17 selected characteristics) less likely to spend more on health than NRRH. RRH were more likely to spend more on health in those households where there where boys, children between 11-17 years, elderly males and a presence of biological father for all children. These results could raise questions of whether there may be a possible presence of gender preference. The age group 11-17 years fall in the mandatory school enrolment requirement age group in Jamaica (Ministry of Justice Jamaica, 2014). Given that the results of previous studies suggest that remittance receipts in Jamaica are generally used to cover basic expenses to include bills, educational and health cost, this result may be an indication of the evidence of this situation in the matched dataset.

In the case of expenditure on education, RRH were more likely to spend more than NRRH (13 of 17 selected characteristics). The situation was the reverse of that observed for expenditure on health. NRRH were more likely to spend more on education in those households where there where boys, children between 11-17 years, elderly males and a presence of biological father for all children. Again we see a possible influence of gender relations.

Table 13 Analysis of expenditure in RRH and NRRH based on general household characteristics

HEALTH	RRH		NRRH		Difference
	Mean	Std. Dev	Mean	Std. Dev	
HOUSEHOLD CHARACTERISTICS					
Household Size >3	3,533.18	4,693.66	4,708.77	8,159.43	-1,175.59
All children in the household have their biological father present	5,573.65	13,847.31	5,176.65	9,367.66	397.00
All children in the household have their biological mother present	4,517.46	10,005.44	5,009.29	9,206.14	-491.83
Number of children > 0	5,070.44	10,960.99	5,283.46	9,078.18	-213.02
# of males under 18 years > 0	5,603.51	12,978.67	5,523.90	9,782.08	79.61
# of females under 18 years >0	7,673.33	13,704.09	8,235.67	15,635.59	-562.34
Number of children under 6 yrs > 0	3,937.16	6,492.25	4,375.93	8,263.04	-438.77
Number of children 6-10 yrs > 0	3,548.25	5,119.69	4,099.36	7,080.04	-551.10
Number of children 11-17 yrs > 0	5,820.94	13,475.30	4,698.95	8,427.02	1,121.98
# of males below 15 years in household > 0	5,103.31	10,940.27	5,455.51	10,173.44	-352.20
# of females below 15 years in household	3,473.91	4,583.72	3,814.86	5,697.89	-340.95
# of individuals age 15-64 > 0	6,978.80	13,084.30	7,737.93	17,010.58	-759.14
# of females in working age group (15-64) > 0	6,929.97	12,770.26	7,541.46	14,786.50	-611.49
# of males in working age group (15-64) > 0	6,774.22	13,853.23	7,884.94	18,723.68	-1,110.73
# of males 65+ years in household > 0	11,383.32	29,235.82	11,138.78	15,565.08	244.53
# of females 65+ years in household > 0	10,161.08	15,250.21	11,106.57	17,000.23	-945.50
# of persons in dependent age group in household>0	6,900.69	17,135.32	7,281.05	12,979.80	-380.36

EDUCATION	RRH		NRRH		Difference
	Mean	Std. Dev	Mean	Std. Dev	
HOUSEHOLD CHARACTERISTICS					
Household Size >3	67,106.74	53,489.84	64,625.76	56,321.88	2,480.98
All children in the household have their biological father present	73,437.53	69,656.53	75,311.60	71,581.22	-1,874.07
All children in the household have their biological mother present	64,456.76	61,709.98	61,983.49	58,290.10	2,473.27
Number of children > 0	72,147.15	57,056.88	70,705.25	59,436.77	1,441.90
# of males under 18 years > 0	69,657.72	45,135.03	71,520.29	56,433.90	-1,862.57
# of females under 18 years >0	47,403.85	59,312.12	41,872.03	58,070.63	5,531.82
Number of children under 6 yrs > 0	51,473.77	45,646.85	45,335.14	36,496.30	6,138.63
Number of children 6-10 yrs > 0	75,048.52	56,099.16	71,408.49	52,042.56	3,640.02
Number of children 11-17 yrs > 0	83,718.32	48,272.75	84,810.52	61,484.00	-1,092.20
# of males below 15 years in household > 0	67,626.46	43,616.13	65,254.52	52,077.29	2,371.95
# of females below 15 years in household	70,077.31	61,095.24	61,546.61	49,995.65	8,530.69
# of individuals age 15-64 > 0	39,460.76	56,641.71	36,479.23	55,822.01	2,981.53
# of females in working age group (15-64) > 0	51,908.45	59,750.25	46,452.67	58,631.41	5,455.78
# of males in working age group (15-64) > 0	36,632.81	54,860.52	36,169.28	58,230.30	463.53
# of males 65+ years in household > 0	16,976.49	36,215.74	22,335.66	51,696.24	-5,359.17
# of females 65+ years in household > 0	25,373.66	53,643.49	20,439.74	45,605.13	4,933.92
# of persons in dependent age group in household>0	51,870.64	57,366.03	47,418.59	56,317.22	4,452.05

* 10 %; ** 5 %; *** 1 %

5.1.5 Matched Data: Characteristics of Household Head

When the characteristics of the household heads were considered in this study, statistically significant differences between the households were observed for the expenditure on education. No statistically significant difference was observed for expenditure on health. RRH with female heads, heads with years of schooling in excess of 12 years, and heads who reported in the clerks occupational category were more likely to spend more on education relative to their matched counterparts in NRRH. Further examination of the Occupational categories revealed that RRH heads with stereotypical 'higher level' occupations were more likely to spend more on education than NRRH. The magnitude of the per capita difference in spend also increased with the increase in occupational category. These may be indicative of the presence of tendency to and/or appreciation on the benefits of investment in education human capital. RRH heads who reported at the lower end of the occupational categories were generally found to be less likely to have higher educational expenditures than NRRH. Albeit these results were not found to be statistically significant.

Table 14 Analysis of expenditure in RRH and NRRH based on characteristics of the household head

HEALTH	RRH		NRRH		Difference
	Mean	Std. Dev	Mean	Std. Dev	
HOUSEHOLD HEAD					
Male headed household	7,760.36	17,972.62	8,362.38	17,943.78	-602.02
Female headed household	8,086.57	14,311.34	7,992.29	16,644.17	94.28
Age of household head >64	10,527.04	24,265.67	11,117.70	17,693.09	-590.66
Married household head	11,608.68	25,320.73	9,820.45	15,171.23	1,788.23
Years of schooling of houshold head >12	8,908.58	16,892.21	8,451.71	15,525.25	456.86
Household head - no occupation	9,142.74	12,816.20	8,465.98	16,495.86	676.77
Household head Legislators, Senior Officials and Managers	9,973.59	16,480.65	18,726.45	38,748.21	-8,752.85
Household head Professionals	14,218.49	23,559.48	14,429.74	15,084.71	-211.25
Household head Technicians and Associate Professionals	7,658.36	8,206.31	15,188.20	32,633.40	-7,529.84
Household head Clerks	11,292.49	19,955.74	7,430.52	8,806.91	3,861.97
Household head Service Workers and Shop and Market Sales Workers	4,922.59	7,630.07	4,522.99	7,482.06	399.61
Household head Skilled Agricultural and Fishery Workers	6,600.82	11,889.86	5,191.37	11,458.77	1,409.45
Household head Craft and Related Trades Workers	5,113.26	7,707.88	6,425.13	14,829.85	-1,311.87
Household head Plant and Machine Operators and Assemblers	17,056.56	48,356.11	11,562.45	19,194.66	5,494.11
Household head Elementary Occupations	5,892.77	14,364.12	5,653.65	8,464.55	239.13

EDUCATION	RRH		NRRH		Difference
	Mean	Std. Dev	Mean	Std. Dev	
HOUSEHOLD HEAD					
Male headed household	29,582.56	53,318.47	30,554.06	57,196.68	-971.50
Female headed household	43,606.54	56,347.10	34,973.77	49,188.24	8632.77*
Age of household head >64	16,143.54	33,798.95	18,248.61	42,330.96	-2,105.06
Married household head	49,254.33	65,068.27	42,524.60	66,549.36	6,729.73
Years of schooling of houshold head >12	50,731.14	67,537.72	34,385.64	57,298.16	16345.50***
Household head - no occupation	17,283.19	35,174.65	14,645.86	30,241.52	2,637.33
Household head Legislators, Senior Officials and Managers	64,007.55	73,532.29	51,674.13	71,505.99	12,333.42
Household head Professionals	83,031.19	94,768.03	58,409.38	98,801.37	24,621.82
Household head Technicians and Associate Professionals	77,848.17	79,161.29	50,645.08	65,418.86	27,203.09
Household head Clerks	50,687.99	57,238.33	20,990.20	38,012.66	29697.79**
Household head Service Workers and Shop and Market Sales Workers	34,077.52	42,593.03	36,922.58	46,994.48	-2,845.06
Household head Skilled Agricultural and Fishery Workers	20,097.99	43,856.32	30,972.83	54,735.35	-10,874.84
Household head Craft and Related Trades Workers	31,244.50	52,018.88	30,401.14	42,764.42	843.36
Household head Plant and Machine Operators and Assemblers	41,969.99	41,124.27	42,075.37	60,757.71	-105.39
Household head Elementary Occupations	30,639.20	41,556.13	29,374.77	41,722.17	1,264.43

* 10 %; ** 5 %; *** 1 %

5.1.6 Matched Data: Characteristics of Location

Geographical location is an important characteristic to consider as where people live may introduce varied dynamics in their expenditure patterns. Examining location of households and their expenditure on health we observed no significant difference in spend between RRH and NRRH in rural compared to urban locations. However when counties were observed, RRH in Surrey county were found to have significantly smaller health expenditure than NRRH.

In terms of education, RRH in urban areas were found to be more likely to have higher educational expenditure relative to NRRH. This could result from a number of factors including but not limited to the fact that schools in the urban areas may be more expensive than rural areas, more attention placed on the relevance of human capital in the urban area, more robust regulatory infrastructure in urban areas (specifically in cases of mandatory schooling) or higher incomes.

Table 15 Analysis of expenditure in RRH and NRRH based on geographical location of household

HEALTH	RRH		NRRH		Difference	EDUCATION	RRH		NRRH		Difference
	Mean	Std. Dev	Mean	Std. Dev			Mean	Std. Dev	Mean	Std. Dev	
LOCATION						*LOCATION*					
Urban household	9,442.25	20,403.09	9,638.48	21,058.98	-196.23	Urban household	43,489.83	59,479.06	32,521.26	55,144.02	10968.57**
Rural household	6,191.89	10,216.91	6,721.89	12,365.25	-530.00	Rural household	27,093.76	48,348.26	32,458.67	52,552.32	-5,364.92
Cornwall County	7,649.11	11,521.10	5,646.26	8,644.50	2,002.85	Cornwall County	32,200.12	47,165.06	29,530.04	49,488.74	2,670.08
Middlesex County	8,967.74	20,292.07	8,216.54	18,772.21	751.20	Middlesex County	32,654.71	50,117.40	31,563.54	49,422.60	1,091.17
Surrey County	6,465.11	13,076.33	10,576.17	20,440.35	-4111.06**	Surrey County	43,619.04	67,108.51	36,921.87	64,358.40	6,697.17

* 10 %; ** 5 %; *** 1 %

5.1.7 Matched Data: Other Socio-Economic Characteristics

A significant difference was found in health expenditure favoring NRRH. All other observed characteristics in this group were found to have insignificant differences for health expenditure across the two household groups in this study. While these other characteristics were insignificant, it was generally viewed that RRH were less likely to record more health expenditure than NRRH. One surprising find was that there was a significant difference between RRH and NRRH's health expenditure when at least one of its members had a chronic illness. NRRH households were more likely to spend more than RRH.

Two characteristics were found to be significant when the difference in educational expenditure was assessed for RRH and NRRH. RRH with at least one member with health insurance and households scoring combined health status of 3/5 (estimated, see Appendix G, p70) were found to have higher expenditure on education than NRRH. One would have expected such an observation or similar observation for the expenditure on health. Besides the limitations highlighted later in this study it is not clear why this result was obtained. As such this study's result diverts from Diether Beuermann, Ruprah & Ricardo Sierra, 2014 and Valero-Gil, 2008 who found that a significant differences in expenditure on insurance and remittances based on the type of household (RRH, NRRH).

Table 16Analysis of expenditure in RRH and NRRH based on other socio-economic characteristics

HEALTH	RRH		NRRH		Difference
	Mean	Std. Dev	Mean	Std. Dev	
OTHER SOCIO-ECONOMIC					
Per Capita Population Quintile 1	2,105.92	4,391.95	1,968.33	2,954.08	137.59
Per Capita Population Quintile 2	3,402.32	4,054.07	3,550.25	5,192.54	-147.94
Per Capita Population Quintile 3	6,064.51	11,644.48	6,955.92	13,614.91	-891.41
Per Capita Population Quintile 4	5,940.58	7,569.54	6,202.04	8,121.69	-261.46
Per Capita Population Quintile 5	14,072.65	24,679.60	15,024.38	26,606.63	-951.72
Per Capita Population Quintile >2	9,849.67	18,849.15	10,218.91	19,784.47	-369.24
# of household members in school > 0	5,339.64	10,937.41	5,792.98	9,413.60	-453.34
# of household members in pre-primary school > 0	4,078.64	6,715.35	4,352.80	8,437.67	-274.16
# of household members in primary school > 0	4,281.23	9,768.11	4,209.03	6,748.65	72.20
# of household members in secondary school > 0	5,631.65	11,694.21	4,912.37	8,356.16	719.28
# of household members in technical or vocational school > 0	6,337.22	8,968.16	3,743.75	6,272.04	2,593.47
# of household members in tertiary school > 0	9,639.78	19,332.70	8,940.36	10,615.29	699.42
# of household members in other types of school > 0	6,911.63	9,299.36	11,123.52	11,099.40	-4,211.89
Share of household members with health insurance>0	12,201.43	18,350.71	14,794.36	22,189.55	-2,592.94
Number of chronic illnesses per houshold member>0	8,311.36	12,933.83	11,356.89	17,488.30	-3045.52*
general health score of the household 1	5,505.88	12,804.46	8,212.50	20,255.59	-2,706.63
general health score of the household 2	7,660.51	14,523.45	7,390.85	20,593.31	269.67
general health score of the household 3	16,010.02	33,692.48	13,464.85	17,206.02	2,545.18
general health score of the household 4	11,977.78	12,243.96	8,361.11	9,601.88	3,616.67
general health score of the household 5	29,000.00	0.00	100.00	0.00	28,900.00
general health score of the household >2.5	14,300.96	29,017.41	10,995.98	15,278.60	3,304.98

EDUCATION	RRH		NRRH		Difference
	Mean	Std. Dev	Mean	Std. Dev	
OTHER SOCIO-ECONOMIC					
Per Capita Population Quintile 1	26,614.05	30,172.11	27,818.25	35,318.26	-1,204.20
Per Capita Population Quintile 2	36,285.94	37,055.71	35,709.82	38,672.92	576.12
Per Capita Population Quintile 3	38,075.48	50,457.18	37,511.21	41,719.77	564.26
Per Capita Population Quintile 4	47,005.76	63,550.06	39,152.35	64,988.30	7,853.41
Per Capita Population Quintile 5	30,849.41	64,403.77	25,548.11	62,665.01	5,301.30
Per Capita Population Quintile >2	37,038.79	61,145.10	32,832.15	58,860.89	4,206.64
# of household members in school	71,593.98	59,226.39	68,056.02	60,431.25	3,537.96
# of household members in pre-primary school	51,727.08	45,703.40	46,114.01	33,841.25	5,613.08
# of household members in primary school	75,698.88	53,883.37	74,352.33	59,368.52	1,346.55
# of household members in secondary school	98,291.30	50,313.01	93,533.31	56,652.62	4,757.99
# of household members in technical or vocational school	88,321.28	39,260.48	96,721.88	36,906.65	-8,400.60
# of household members in tertiary school	40,340.34	44,295.06	26,840.54	36,146.31	13,499.80
# of household members in other types of school	13,642.05	20,340.21	25,263.33	45,526.86	-11,621.28
Share of household members with health insurance	56,912.41	72,706.89	39,160.59	63,181.43	17751.82**
Number of chronic illnesses per houshold member	39,294.52	54,551.68	33,676.59	52,374.80	5,617.93
general health score of the household 1	34,189.99	55,944.47	35,790.02	62,406.49	-1,600.03
general health score of the household 2	31,612.11	53,323.52	24,021.77	41,572.59	7,590.34
general health score of the household 3	9,636.17	34,220.11	1,072.32	8,024.52	8563.85*
general health score of the household 4	0.00	0.00	0.00	0.00	0.00
general health score of the household 5	0.00	0.00	0.00	0.00	0.00
general health score of the household >2.5	7,725.03	29,351.78	4,150.08	15,852.93	3,574.95

5.2 Sensitivity Analysis

A good way of testing the robustness of the nearest neighbor matching approach could be to employ another matching method and assess the results. A first choice would be to use unrestricted greedy matching. However this approach requires that the number of cases in the control group (NRRH) exceed those in the treated group (RRH). In this study this did not hold. As such robustness checks were done by employing nearest neighbor matching with replacement using caliper of 0.005. One-to-many matching with a ratio of 5 was used. This meant that up to five control cases (NRRH) could be matched with a treated case (RRH).

With this refined approach improvement in the results matching outcome would be expected in terms of increased number of matches. On the other hand, this approach could also affect the balance in covariates and may cause imbalance. So in effect we could increase the matches at the expense of reducing the power of the results. Note that the same caliper (0.005) was maintained. This was to ensure that the cases matched stayed within the area of common support. Though some differences in the result would be anticipated, the major findings observed in the analysis of the expenditure outcomes would be expected to be similar.

5.2.1 Results of Matching Using Replacement

Table 17 Sample sizes of RRH and NRRH post matching with replacement exercise

	All		Matched		Unmatched		Discarded	
Subsamples	Control	Treated	Control	Treated	Control	Treated	Control	Treated
(all cases)	812	888	696	725	116	163	0	0

As expected the number of matches increased using matching with replacement. The number of household in the new matched dataset increased by 403 to 1421.

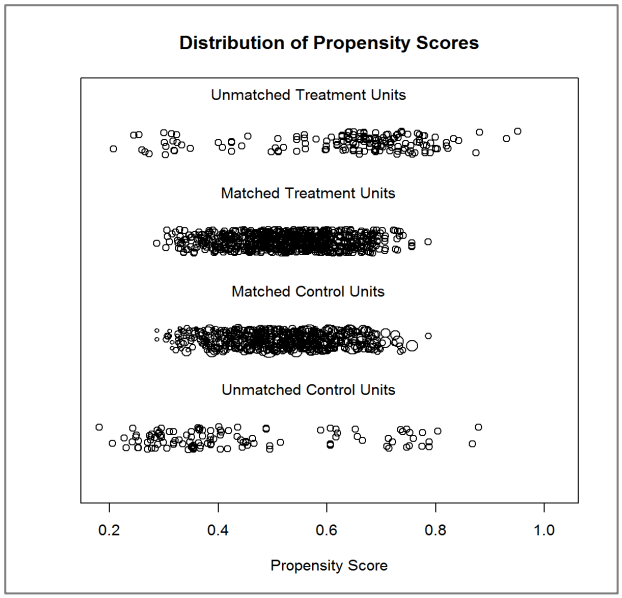
Table 18 Relative multivariate imbalance $\mathcal{L}_1$ (Iacus, King, & Porro, 2010)

	Before Matching	After Matching
(all cases)	.999	.999

The $\mathcal{L}_1$ post matching turned out to be the same as $\mathcal{L}_1$ pre- matching. Ideally a smaller $\mathcal{L}_1$ would be required to suggest that matching improved the overall balance imbalance in the covariates. However other assessments suggest that the results may be favorable in some respect.

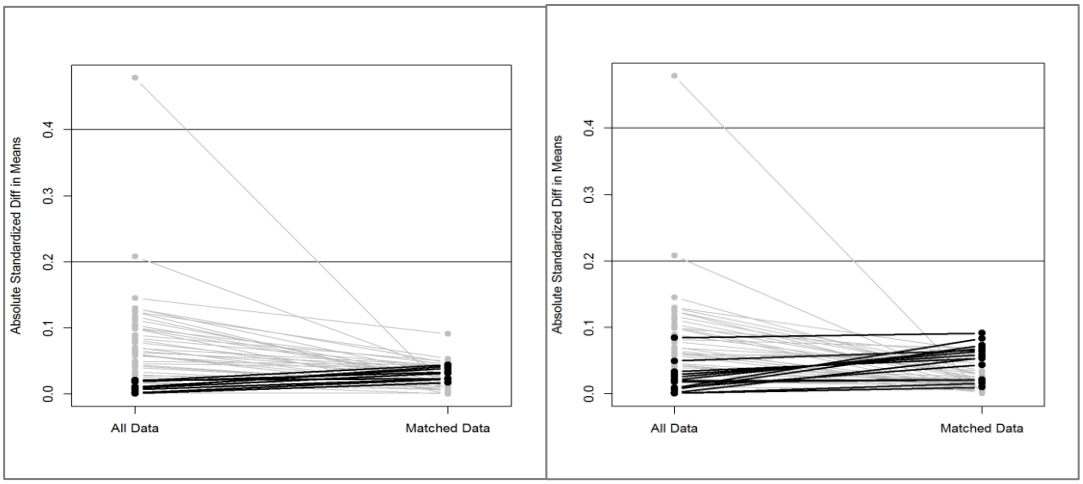
A brief look at the dot plot of the distribution of the propensity scores suggests that the spread of the outliers were reduced in the matched dataset. This is evidenced by the reduction in the number of clear dots in the matched groups.

Figure 9 Dot plot of the distribution of the propensity scores after matching with replacement



The figures below show the line plot of standardized differences in the means before and after matching using for the first matched dataset without replacement and the new dataset with replacement. Recall that best results are achieved when the standardized differences tend towards 0 and the dark lines are minimized. Matching with replacement did not improve on the previous dataset generated. However note that that all standardized differences were below 0.10 despite the presence of increased standardized differences post the matching exercise.

Figure 10 Line plot of standardized differences in the means for matched samples without and with replacement



Similar results were found looking at the histogram to examining the distribution of the propensity scores for both groups in both datasets. While the initial dataset seemed to be smoother it was found that the tails of the distribution in both datasets of the propensity scores narrowed post matching resulting in a more compressed distribution. Albeit in the second instance (matching with replacement) seems to have a slightly less favourable distribution.

Figure 11 Histogram of the distribution of the propensity scores in matched datasets without and with replacement

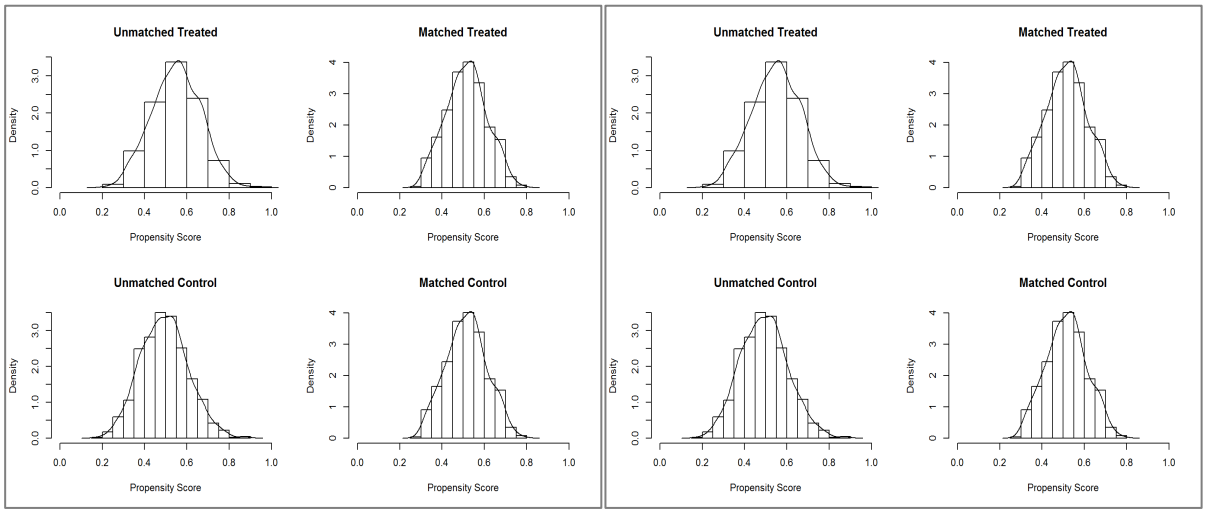
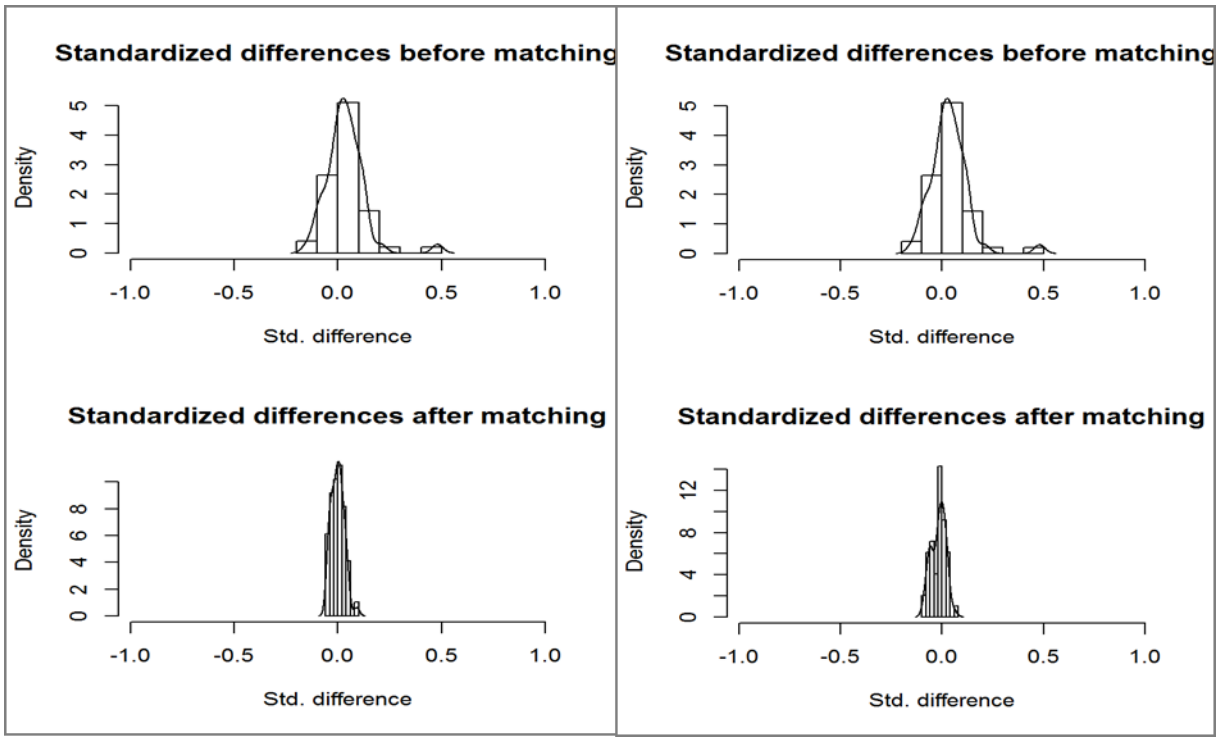


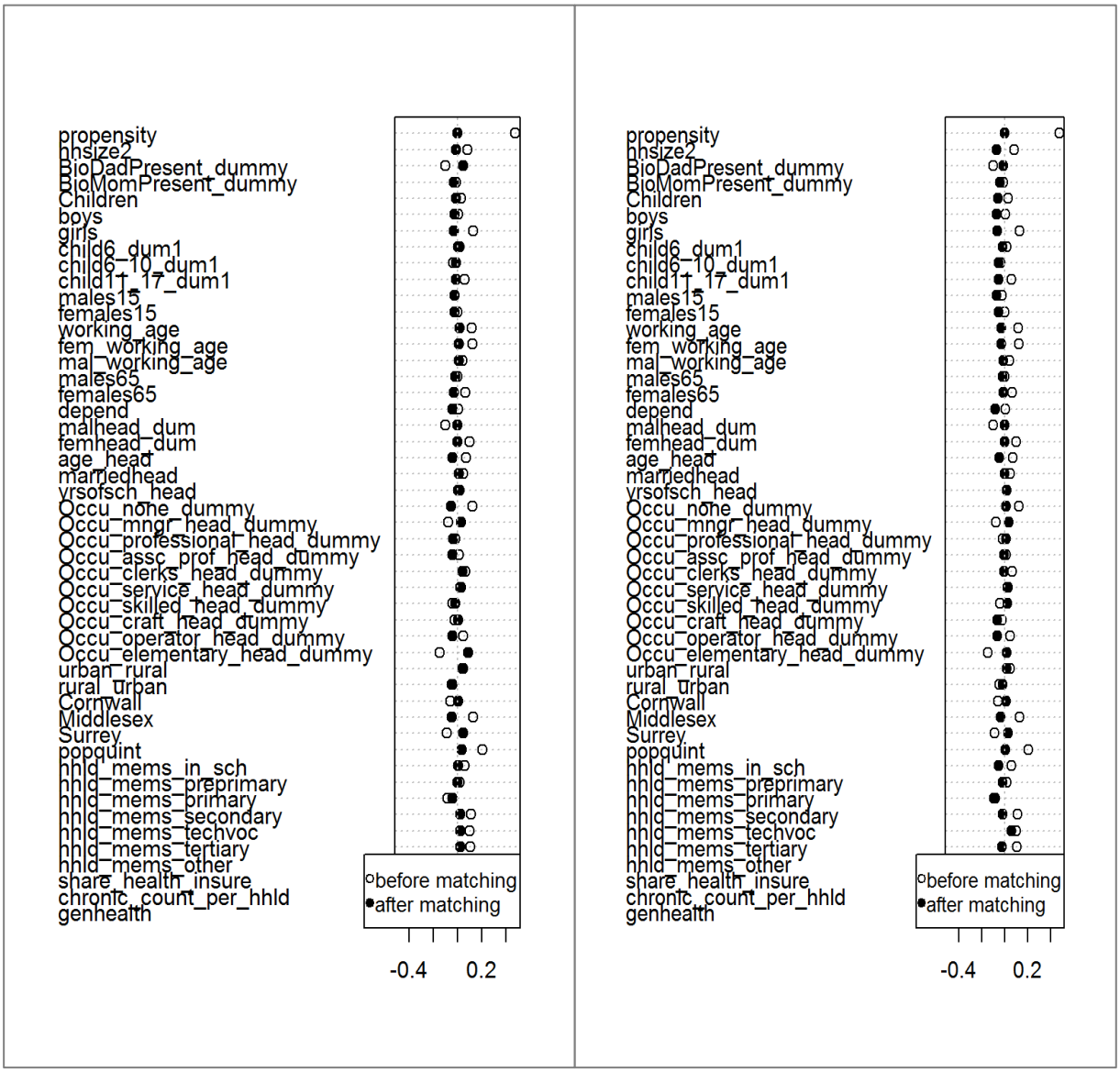
Figure 12 plot of standardized differences using matching without and with replacement



The plot of the standardized differences before and after matching also show a minimization of the differences tending towards zero. The results of both matching methods were relatively similar.

Finally, the dot plot of individual propensity scores for both datasets again highlights the similarity in the outcomes of the balancing analysis of both matching methods, though matching with replacement seem to have more imbalance than without replacement.

Figure 13 Dot plot of individual propensity scores using matching without and with replacement



5.2.2 Analysis of New Dataset (Matching with Replacement)

The final result was similar to matching without replacement when the differences in the expenditure between the two household types were examined using the new matched dataset. There was no significant difference in the expenditure on health nor education between RRH and NRRH. Though similar in terms of no significant differences, it was observed that the direction of the relationship changed for the difference in expenditure on health. In both cases the difference in expenditure for RRH was positive. In the previous case it was only positive for differences in expenditure on education.

Table 19 Test for differences in expenditure between RRH and NRRH (matched sample with replacement)

	RRH		NRRH		Difference
	Mean	**Std. Dev**	**Mean**	**Std. Dev**	
Expenditure					
Health	8 699.25	17 795.20	8 111.35	19 151.44	587.90
Education	34 170.95	52 735.75	32 003.73	52 209.19	2 167.22

Further analysis of the three major categories used in this study to examine the expenditure between RRH and NRRH revealed that there were more significant differences observed for both categories of expenditure in the matched sample with replacement. The following tables highlights the characteristics that yielded significant differences in expenditure on health between RRH and NRRH in the first instance, and on education in the second instance. The highlighted variables are those that were found to be significant at least in the initial matching without replacement. The complete table with all variables may be found in Appendix E. These results may suggest some improvement as there were significant differences observed in all three major categories identified. This improvement may be as a result of the relatively relaxed nature of matching with replacement. It affords repeated selection of best matches. The drawback however is that the power of the results is reduced. The results from the sensitivity analysis also concurs with the literature on propensity score matching that one of the issues with matching approaches, much like other approaches, is that there is no one fixed way of doing it and the approach may lead to completely different results (Fullerton et al., 2016).

Table 20 Analysis of expenditure on health between RRH and NRRH based on major categories

HEALTH	RRH		NRRH		Difference
	Mean	Std. Dev	Mean	Std. Dev	
HOUSEHOLD CHARACTERISTICS					
Number of children 11-17 yrs > 0	6 423.57	15 211.17	4 446.54	7 622.02	1977.03**
# of females below 15 years in household	5 372.77	12 891.40	3 911.68	5 788.20	1461.09**
HOUSEHOLD HEAD					
Female headed household	8 723.18	16 504.24	8 312.43	19 666.33	410.75
Household head - no occupation	9 716.06	17 631.40	7 132.64	14 610.72	2583.42*
Household head Professionals	20 647.22	33 711.25	14 596.02	14 341.81	6051.20*
Household head Technicians and Associate Professionals	6 348.32	7 717.45	24 617.21	47 338.91	-18268.88**
Household head Clerks	10 901.33	17 425.45	6 372.68	7 823.40	4528.65**
Household head Service Workers and Shop and Market Sales Workers	5 761.04	9 879.86	3 614.96	6 011.11	2146.08***
LOCATION					
Cornwall County	8 834.50	16 567.65	5 392.93	8 674.36	3441.58***
Surrey County	7 436.72	15 772.25	11 260.59	24 401.25	-3823.86**
OTHER SOCIO-ECONOMIC					
# of household members in school	6 568.98	14 386.62	5 479.07	9 181.68	1089.91*
# of household members in primary school	4 734.17	9 654.92	4 318.84	6 621.73	415.33**
Number of chronic illnesses per houshold member	10 601.15	15 498.37	10 270.39	16 303.85	330.76
general health score of the household 1	5 029.18	11 309.75	9 438.75	25 845.78	-4409.57*
general health score of the household 4	24 635.71	41 668.02	7 193.06	7 578.62	17442.66***
general health score of the household >2.5	16 394.96	30 588.08	10 703.69	14 485.33	5691.27***

* 10 %; ** 5 %; *** 1 %

Table 21Analysis of expenditure on education between RRH and NRRH based on major categories

EDUCATION	RRH		NRRH		Difference
	Mean	Std. Dev	Mean	Std. Dev	
HOUSEHOLD CHARACTERISTICS					
Number of children under 6 yrs > 0	48 454.12	44 154.55	43 105.31	33 576.50	5348.80*
# of females below 15 years in household	66 929.91	57 258.05	59 158.39	43 719.85	7771.51**
HOUSEHOLD HEAD					
Female headed household	41 121.46	54 016.59	36 078.69	50 440.34	5 042.78
Years of schooling of houshold head >12	45 418.65	63 596.44	34 466.58	57 059.31	10952.07***
Household head Clerks	47 948.56	53 545.21	24 116.15	39 560.39	23832.41***
LOCATION					
Urban household	39 290.61	57 184.60	32 869.68	55 128.30	6420.92*
OTHER SOCIO-ECONOMIC					
Share of household members with health insurance	50 374.45	68 226.32	35 745.48	59 556.32	14628.97***
general health score of the household 2	31 973.31	52 731.06	23 179.25	40 647.21	8794.06***
general health score of the household 3	7 240.91	29 193.13	1 968.85	10 722.94	5272.06**

* 10 %; ** 5 %; *** 1 %

5.3 Limitations

5.3.1 Propensity Score

Since the true propensity score is unknown in non-experimental studies, one can never a 100% confident that the estimates are accurate. Researchers have cautioned against this limitation in the use of propensity score analysis for these studies. (King & Nielsen, 2018) asserts that that these scores should not be used for matching. To guard against this limitation it is important that the iterative process of checking the propensity score for balance be exercised with much care (Rosenbaum & Rubin, 1983). An alternative to iterative checks is Genetic matching which uses a search algorithm to determine the weight of each covariate (Diamond & Sekhon, 2013).

The results of propensity score matching methods are not easy to replicate. Researchers using PSM may achieve different results owing to the fact that there are many ways to approach the matching (Fullerton et al., 2016) and in some cases some decisions are made that may be preferential.

5.3.2 Effect of Household Characteristics

Household size could possibly have either a negative or positive impact on remittance receipts and thus affect expenditure. This may be influenced by economies of scale in consumption, the rate of decline in marginal utility of home consumption and the existence of preference for a subset of the receiving household by the remitter (Samuel, 2004).

5.3.3 Omitted Variable Bias

Omitted variable bias may still be an issue. Matching does not solve the issue of unobserved confounders. It is therefore important to acknowledge that the observables may differ in other unobserved ways that may not have been accounted for in this study. For example two households matching perfectly based on their propensity scores but they in fact were two different types (RRH vs NRRH) then there is likely some other characteristic(s) that would make the household fall in one of the two groups – an unobserved confounder(s).

5.3.4 Remittances in Kind

Although there was a measure of remittances in kind in the JSLC, it was not included in this study because of the increased recall bias and questions of reliability in quantifying these remittances. This omission needs to be taken into consideration when using the results.

5.3.5 Survey Data

Survey data may suffer from recall bias. Questions in the JSLC reference the last year. Therefore respondents were required to recall events that took place a year ago. Providing correct responses especially ones relating to finances may be very difficult unless these were being recorded in a detailed way throughout the year. Though this might exist for 'big spends' or for those who keep detailed household budgets. This may not be common place among many households. Also generating annualize data from monthly data in surveys can be problematic and sometimes provide inaccurate estimates.

5.3.6 Area of Common Support

To improve the balance on the covariates, units outside the area of common support were discarded from the sample. As a result of this the estimate of interest changes and the causal effect for units are localized to the sample of comparable units. The 509 households with matching characteristics as per the restrictions of this study.

5.3.7 Scope of Remittances Receipts

In this study focus was placed on international remittance receipts to households. There exists the possibility that there may be internal transfers to household that may have significant impact on their expenditure patterns which this study would not have examined. The study is also limited by the absence of recent data which specifically speaks to international remittance receipts and how these receipts are spent.

Using data from a topic specific survey on remittances as is done in other countries could possible improve the extent to which empirical analysis on the expenditure patterns of remittance receiving households. In such surveys more specific questions related to the

migration are posed. This would facilitate the ability to offer a deeper look into the dynamics of remittance receipts, migration, and household expenditure.

5.3.8 Characteristics of Household Heads

Household head characteristics such as occupation, age, sex were investigated mostly in this study as it was assumed that characteristics of the household head may influence the expenditure patterns. However, it could be that similar characteristics of other adult members of the household could affect the expenditure patterns especially in cases where the difference in the number of working adults in households is significant.

5.3.9 Size of Control Group

More favourable results could possibly be obtained where the control group has significantly more observations than the treatment group to aid in getting more matches. In this study there were more observations in the treatment group than in the control group.

5.3.10 Magnitude of Remittance Receipt

Recall that this study did not account for the magnitude of the remittance receipt. Households were classified as remittance receiving simply on the fact that they responded positively to at least one of the remittance receiving filters. Use of households with a specified percent share of household income coming from remittance receipts would be a better indication of a remittance receiving household. Especially if we intend to examine the expenditure patterns in light of remittance receipts. This alternative approach was not explored in this study due to data issues due to coding of the currencies in the dataset, the fact that the JSLC was not designed with a target explore remittances or migration in depth along with the possibility of seasonality affecting the reports of remittances.

5.3.11 Variable Creations, Recoding

The interpretation and use of the results rely heavily on the methodology of the variable creation process. The denominator for the per capita education expenditure variable was the number of persons attending school during the survey period. While this may be an ideal/recommended measure, it may have its limitations as households may report educational expenditure even if none of their members were attending school during the period. The difference in calculation method for per capita expenditure for health and education impacts how we interpret and to what extent we compare the two outcomes since all household members were used in calculating health versus only those who reported educational activity were included in the calculation of per capita expenditure.

5.3.12 Remitter Characteristics

The characteristics of the remitter could also influence how household budget is spent especially if the remitter was once the head of the sending household. This was not accounted for on this study.

6 Conclusion

Using data extracted from the 2015 round of the Jamaica Survey of Living Conditions, this study sought to assess whether remittance receiving households' expenditure on health and education differed from non-remittance receiving households by employing propensity score matching. The findings suggest that there was no statistically significant difference in these households' expenditure on health for the households observed in the study. There was also no statistically significant difference in their expenditure on education. Households in this study spent more on education versus health. This result held true for both RRH and NRRH suggesting that there may exist a relatively higher importance placed on education.

Selected characteristics of the household heads seem to influence the expenditure on education. RRH households with female heads and household heads with years of schooling in excess of 12 years showed significant positive differences in their expenditure on education. These same characteristics considered, there were no significant difference observed in the expenditure on health. RRH having heads with stereotypically 'higher level' occupations were more likely to spend more on education than those in NRRH. Albeit this result was insignificant. NRRH households were more likely to spend more than RRH on health when at least one member had a chronic illness.

Location also mattered. RRH in urban areas were found to be more likely to have higher educational expenditure relative to NRRH. NRRH in the county of Surrey were more likely to spend on more on health than RRH in that county.

There seemed to be a relationship between health expenditure on education. RRH with at least one member with health insurance and households scoring combined health status of 3/5 were found to have higher expenditure on education than NRRH.

The general differences observed in the dataset prior to matching were relatively small and this could have contributed to the insignificant results that were found in this study. This could have also been exacerbated by the absence of a remittance specific survey. This highlights the effect that data availability can have on the outcome of research of this nature. The insignificant results found is not sufficient to assert that there is no difference between the expenditure on health and education in remittance receiving versus non-remittance receiving households in Jamaica. Generalizations about Jamaica cannot be made. This insignificance only relate to the

1018 households that were observed. The extent to which remittance receipt would influence household expenditure is also linked to the share of household income that is received from overseas. Since this was not observed in this study, the interpretation and use of these results are cautioned.

Strength of generalization from observational studies rely heavily on the availability of good data. Numerous studies before in other jurisdictions have found significant relationships using data that was generated from surveys with special focus on remittance receipts and expenditure of those receipts. It may therefore be necessary that the Jamaican government invest in such an undertaking on a regular basis to improve the quality and availability of data on remittance receipts given the relative magnitude of these receipts and its economic benefit to the country. Once this is in place the strength of future research is likely to improve and provide more insight on the impact of the remittance receipts.

References

Adams, R. H. & Page, J. (2005). Do International Migration and Remittances Reduce Poverty in Developing Countries?, *World Development*, vol. 33, no. 10, pp.1645–1669.

Alejandra, C. & Manuelita, U. (2003). INTERNATIONAL MIGRATION, REMITTANCES, AND SCHOOLING: EVIDENCE FROM EL SALVADO, in *NATIONAL BUREAU OF ECONOMIC RESEARCH*, June 2003.

Amuedo-Dorantes, C. & Pozo, S. (2011). New Evidence on the Role of Remittances on Healthcare Expenditures by Mexican Households, *Review of Economics of the Household*, vol. 9, no. 1, pp.69–98.

Antón, J.-I. (2010). The Impact of Remittances on Nutritional Status of Children in Ecuador, *The International Migration Review*, vol. 44, no. 2, pp.269–299.

Arif, I., Raza, S. A., Ivkovi, A. F. & Suleman, T. (2018). The Role of Remittances in the Development of Higher Education: Evidence from Top Remittance Receiving Countries, *Social Indicator Research*.

Asian Development Bank & World Bank. (2018). Migration and Remittances for Development in Asia, Manila, Philippines: Asian Development Bank, Available Online: https://www.adb.org/publications/migration-remittances-developing-asia [Accessed 4 August 2019].

Bank of Jamaica. (2018a). Remittance Bulletin Jamaica, December 2018, Jamaica: Central Bank of Jamaica, p.1, Available Online: http://www.boj.org.jm/uploads/pdf/rem_updates/rem_updates_dec2018.pdf [Accessed 12 May 2019].

Bank of Jamaica. (2018b). Inbound Remittances to Jamaica by Corridor (through Remittance Companies), Jamaica: Central Bank of Jamaica, Available Online: http://boj.org.jm/uploads/excel/Table_45.xls [Accessed 12 May 2019].

Bertsekas, D. P., & Tseng, P. (1988). Relaxation methods for minimum cost ordinary and generalized network flow problems. Operations Research, 36(1), 93-114.

Bertsekas, D. P., & Tseng, P. (1988). Relaxation methods for minimum cost ordinary and generalized network flow problems. Operations Research, 36(1), 93-114.

Borjas, G. J. (1994). The Economics of Immigration, *Journal of Economic Literature*, vol. 32, no. 4, pp.1667–1717.

Chiuzan, C. (2018). Propensity Score Matching, A Practical Tutorial, Biostatistics, Epidemiology and Research Design, Columbia University, 19 March 2018.

Chort, I. & Senne, J.-N. (2015). Selection into Migration within a Household Model: Evidence from Senegal, *The World Bank Economic Review*, vol. 29, no. suppl_1, pp.S247–S256.

Clarke, G. R. G. & Wallsten, S. J. (2003). Do Remittances Act Like Insurance? Evidence From a Natural Disaster in Jamaica, *SSRN Electronic Journal*, [e-journal], Available Online: http://www.ssrn.com/abstract=373480 [Accessed 19 May 2019].

Dehejia, R. H. & Wahba, S. (2012). PROPENSITY SCORE-MATCHING METHODS FOR NONEXPERIMENTAL CAUSAL STUDIES, *THE REVIEW OF ECONOMICS AND STATISTICS*, p.11.

Diamond, A. & Sekhon, J. S. (2013). Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies, *Review of Economics and Statistics*, vol. 95, no. 3, pp.932–945.

Diether Beuermann, Ruprah, I. & Ricardo Sierra. (2014). Remittances as a Safety Net in Jamaica, Policy Brief, IDB-PB-235, Inter American Development Bank, p.13.

Dietz, B., Gatskova, K. & Ivlevs, A. (2015). Emigration, Remittances and the Education of Children Staying Behind: Evidence from Tajikistan, p.23.

Fullerton, B., Pöhlmann, B., Krohn, R., Adams, J. L., Gerlach, F. M. & Erler, A. (2016). The Comparison of Matching Methods Using Different Measures of Balance: Benefits and Risks Exemplified within a Study to Evaluate the Effects of German Disease Management Programs on Long-Term Outcomes of Patients with Type 2 Diabetes, *Health Services Research*, vol. 51, no. 5, pp.1960–1980.

Garrido, M. M., Kelley, A. S., Paris, J., Roza, K., Meier, D. E., Morrison, R. S. & Aldridge, M. D. (2014). Methods for Constructing and Assessing Propensity Scores, *Health Services Research*, vol. 49, no. 5, pp.1701–1720.

Hansen, B. B. (2004), Full Matching in an Observational Study of Coaching for the SAT. Journal of the American Statistical Association, 99, 609-618.

Hansen, B. B., & Klopfer, S. O. (2006). Optimal full matching and related designs via network flows. Journal of Computational and Graphical Statistics, 15(3).

Hansen, B. B., & Bowers, J. (2008). Covariate balance in simple, stratified and clustered comparative studies. Statistical Science, 2, 219-236.

Heinrich, C. J., Maffioli, A. & Vázquez, G. M. (2010). A Primer for Applying Propensity-Score Matching Impact-Evaluation Guidelines, 2010.

Henry, C., Moulton, J. & Ricketts, J. (2009). Motives for Sending Remittances to Jamaica: An Application of the BPM61 Definition of Remittances, *Central Bank of Jamaica*, p.36.

Hines, A. (2014). Migration, Remittances and Human Capital Investment in Kenya, p.35.

Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2011). MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. Journal of Statistical Software, 42(8).

Ho, D., Imai, K., King, G., & Stuart, E. (2007), Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference. Political Analysis, 15, 199-236.

Iacus, S. M., King, G., & Porro, G. (2009). CEM: Software for coarsened exact matching. Journal of Statistical Software, 30, 1-27.

International Monetary Fund (ed.). (2014). BPM6 Compilation Guide, Washington, DC: International Monetary Fund.

Irving, R., Charlton, V., Morrison, E., Facey, A. & Buchanan, O. (2013). Demographic Characteristics of World Class Jamaican Sprinters, *The Scientific World Journal*, vol. 2013, pp.1–5.

King, G. & Nielsen, R. (2018). Why Propensity Scores Should Not Be Used for Matching, *Political Analysis*, pp.1–20.

McKenzie, D. & Rapoport, H. (2011). Can Migration Reduce Educational Attainment? Evidence from Mexico, p.44.

McLean, E. A. (2008). An Investigation of Recent Trends In the Remittance Industry: Evidence from Jamaica, p.36.

Ministry of Justice Jamaica. (2014). Child Care and Protection Act, Revised 2014.

Mora, J. J. (2013). GENDER DIFFERENCES BETWEEN REMITTANCES AND LABOR PARTICIPATION IN DEVELOPING COUNTRIES: A CROSS-SECTION ANALYSIS OF COLOMBIA IN YEAR 2008, vol. 13, p.14.

Pan, W. & Bai, H. (2015). Propensity Score Analysis, Concepts and Issues, p.17.

Ramocan, E. G. (2011). Remittances to Jamaica Findings from a National Survey of Remittance Recipients, Kingston: Bank of Jamaica.

Rapoport, H. & Docquier, F. (2005). The Economics of Migrants' Remittances, p.82.

Ricketts, J. R. (2011). Impact of Macroeconomic Shocks on Remittance Inflows to Jamaica: A VECM Approach., *Central Bank of Jamaica*, p.33.

Rooth, D.-O. & Saarela, J. (2007). Selection in Migration and Return Migration: Evidence from Micro Data, *Economics Letters*, vol. 94, no. 1, pp.90–95.

Rosenbaum, P. R. & Rubin, D. B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects, *Oxford Journal*, vol. 70, no. 1, pp.41–55.

Samuel, W. (2004). Migration and Remittances: A Case Study of the Caribbean.

Shadish, W. R., Clark, M. H. & Steiner, P. M. (2008). Can Nonrandomized Experiments Yield Accurate Answers? A Randomized Experiment Comparing Random and Nonrandom Assignments, *Journal of the American Statistical Association*, vol. 103, no. 484, pp.1334–1344.

Sir Arthur Lewis Institute of Social and Economic Studies (University of the West Indies

(Mona) Jamaica

Stark, O. & Lucas, R. (1985). Motivations to Remit: Evidence from Botswana, *Journal of Political Economy*, vol. 93, no. 5, pp.901–918.

Statistical Institute of Jamaica. (2016). Jamaica Survey of Living Conditions.

Stuart, E. A. (2010). Matching Methods for Causal Inference: A Review and a Look Forward, *Statistical Science*, vol. 25, no. 1, pp.1–21.

Terrelonge, S. C. (2014). For Health, Strength, and Daily Food: The Dual Impact of Remittances and Public Health Expenditure on Household Health Spending and Child Health Outcomes, *Journal of Development Studies*, vol. 50, no. 10, pp.1397–1410.

Thoemmes, F. (2012). PROPENSITY SCORE MATCHING IN SPSS, Available Online: https://arxiv.org/ftp/arxiv/papers/1201/1201.6385.pdf [Accessed 5 May 2019].

Thoemmes, F. J. & Kim, E. S. (2011). A Systematic Review of Propensity Score Methods in the Social Sciences, *Multivariate Behavioral Research*, vol. 46, no. 1, pp.90–118.

Thomas-Hope, E. M. (2002). Caribbean Migration, University of the West Indies Press.

Valero-Gil, J. N. (2008). Remittances and the Household's Expenditures on Health, p.27.

World Bank. (2002). Jamaica Survey of Living Conditions (JSLC) 1988-2000: Basic Information.

World Bank. (2018a). Migration and Remittances: Recent Developments and Outlook, [e-book] World Bank, Available Online: https://www.knomad.org/sites/default/files/2018-12/Migration%20and%20Development%20Brief%2030.pdf [Accessed 2 June 2019].

World Bank. (2018b). Record High Remittances to Low- and Middle-Income Countries in 2017, Available Online: https://www.worldbank.org/en/news/press-release/2018/04/23/record-high-remittances-to-low-and-middle-income-countries-in-2017 [Accessed 6 February 2019].

World CIA Factbook. (2018). Jamaica Economy Profile 2018, Available Online: Jamaica Economy Profile 2018 [Accessed 12 May 2019].

Zachariah, K. C., Mathew, E. T. & Rajan, S. I. (2001). Social, Economic and Demographic Consequences of Migration on Kerala, *International Migration*, vol. 39, no. 2, pp.43–71.

Appendix A: Summary of Dataset Sections used to create required dataset for study

Section Label	Section Code	Type of Data	Number of variables
Health	A	Individual	94
Education	B	Individual	42
Social Protection	D	Individual	15
Daily Expenses	E	Individual	5
Non- Consumption Expenditures	H	Individual	5
Consumption Expenditures	G	Individual	33
Income Received From Sources Outside of Household	K	Individual	46
Annual		Household	33
Poverty Line		Household	46
Principal Earner		Household	10

Appendix B: Section K of questionnaire used to identify remittance receiving households

PART K: MISCELLANEOUS - RECEIVED FROM SOURCES OUTSIDE OF HOUSEHOLD

1 During the past 12 months, has any member of your household received income in cash or in kind from the following sources? PUT A TICK IN THE APPROPRIATE BOX FOR EACH ITEM? ASK QUESTION 1 FOR ALL ITEMS FOR WHICH THE ANSWER IS YES, ASK QUESTION 2.			2 What is the value of the income received by members of your household in cash or in kind from ... [ ] ... during the past 12 months?											
			INDIVIDUAL NUMBER AS IN ROSTER	ITEM CODE	AMOUNT	ORIGINAL CURRENCY	HOW OFTEN IS THIS MONEY / GOODS RECEIVED?		INDIVIDUAL NUMBER AS IN ROSTER	ITEM CODE	AMOUNT	ORIGINAL CURRENCY	HOW OFTEN IS THIS MONEY / GOODS RECEIVED?	
							TIME	PERIOD					TIME	PERIOD
Support for children from parents who live in Jamaica	701	☐ YES ☐ NO												
Support for children from parents who live abroad?	702	☐ YES ☐ NO												
Spouse / Partner who lives in Jamaica	703	☐ YES ☐ NO												
Spouse/ Partner who lives abroad?	704	☐ YES ☐ NO												
Child / children who lives / live in Jamaica	705	☐ YES ☐ NO												
Child / children who lives / live abroad	706	☐ YES ☐ NO												
Other relatives or friends who live in Jamaica	707	☐ YES ☐ NO												
Other relatives or friends who live abroad?	708	☐ YES ☐ NO												
Rental payments for use of land or other property owned by household members?	709	☐ YES ☐ NO												
Social Security (NIS)	710	☐ YES ☐ NO												
Private,Government or other pension fund?	711	☐ YES ☐ NO												
Public Assistance?	712	☐ YES ☐ NO												
Dividend / Interest from loans made by household members or from money deposited in the bank or other financial Institutions?	713	☐ YES ☐ NO												
Windfall receipts ?(lotteries,gambling,inheritances)	714	☐ YES ☐ NO												
Other?	715	☐ YES ☐ NO												

**Daily.....1 Monthly.....4 Yearly.....7
**Weekly.....2 Quarterly.....5 Occassionally.....8
**Fortnightly.....3 Half yearly.....6 Only when requested....9

K

Appendix C: Jamaica's Parishes and Counties

Key

Counties
Cornwall (Hanover, St. Elizabeth, Saint James, and Westmore land)
Middlesex (Clarendon, Manchester, St. Ann, St. Catherine, and St. Mary)
Surrey (Kingston, Portland, St. Andrew, and St. Thomas)

Boundaries
--- Parishes
—— Counties

Miles
0 20

Taken from (Irving et al., 2013)

Appendix D: Matching Iterations with different calipers

***PROPENSITY SCORE MATCHING NEAREST NEIGHBOR WITHOUT REPLACEMENT WITH CALIPER 0F **0.0001** using all 48 VARIABLES**

Sample Sizes

	All		Matched		Unmatched		Discarded	
Subsamples	Control	Treated	Control	Treated	Control	Treated	Control	Treated
(all cases)	812	888	38	38	772	847	2	3

Overall balance test (Hansen & Bowers, 2010)

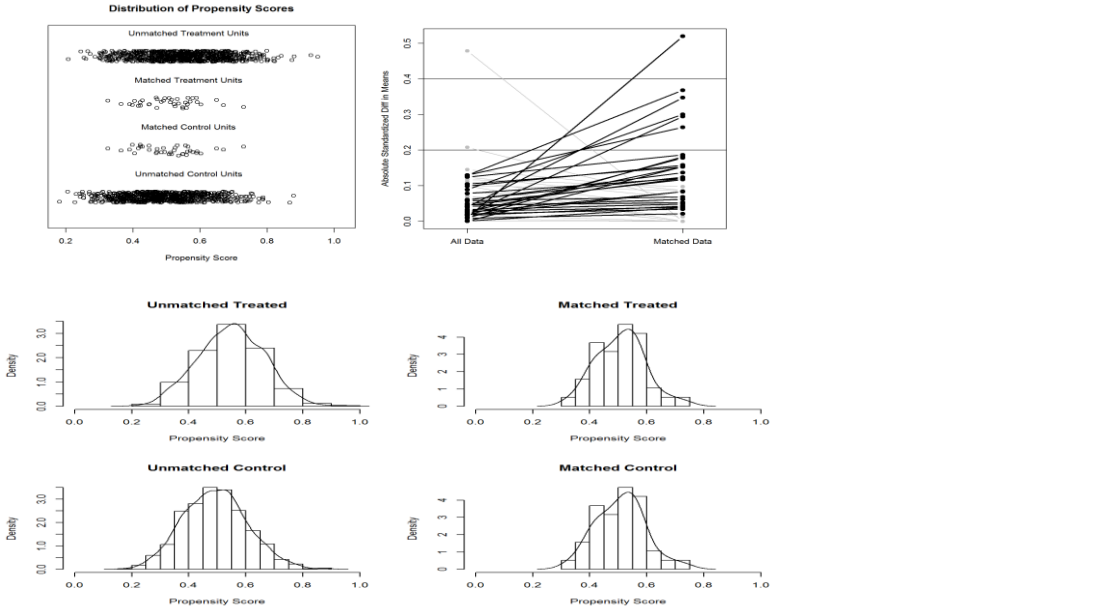
	chisquare	df	p.value
(all cases)	31.611	36.000	.677

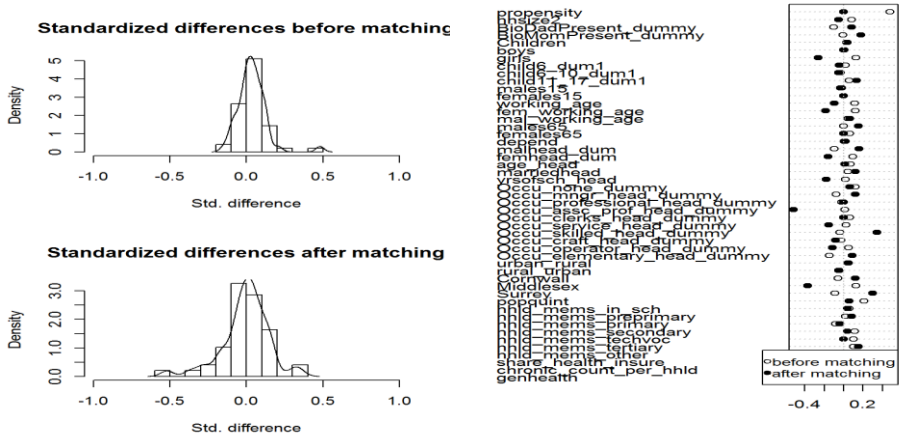
Relative multivariate imbalance L1 (Iacus, King, & Porro, 2010)

	Before matching	After matching
(all cases)	.999	1.000

Summary of unbalanced covariates (|d| > .25)

Subsamples	Covariates	Means Treated	Means Control	SD Control	Std. Mean Diff.
(all cases)	Occu_assc_prof_head_dummy	.000	.105	.311	-.520
	Middlesex	.447	.632	.489	-.368
	Occu_skilled_head_dummy	.289	.158	.370	.347
	Surrey	.289	.158	.370	.301
	hhld_mems_other	.000	.053	.226	-.295
	girls	.868	1.105	.863	-.264





***PROPENSITY SCORE MATCHING NEAREST NEIGHBOR WITHOUT REPLACEMENT WITH CALIPER 0F **0.0075** using all 48 VARIABLES**

Sample Sizes

	All		Matched		Unmatched		Discarded	
Subsamples	Control	Treated	Control	Treated	Control	Treated	Control	Treated
(all cases)	812	888	560	560	250	325	2	3

Overall balance test (Hansen & Bowers, 2010)

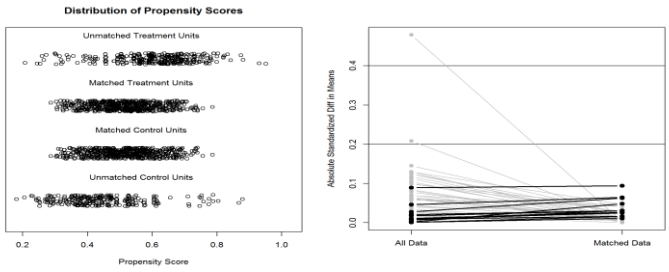
	chisquare	df	p.value
(all cases)	11.308	38.000	1.000

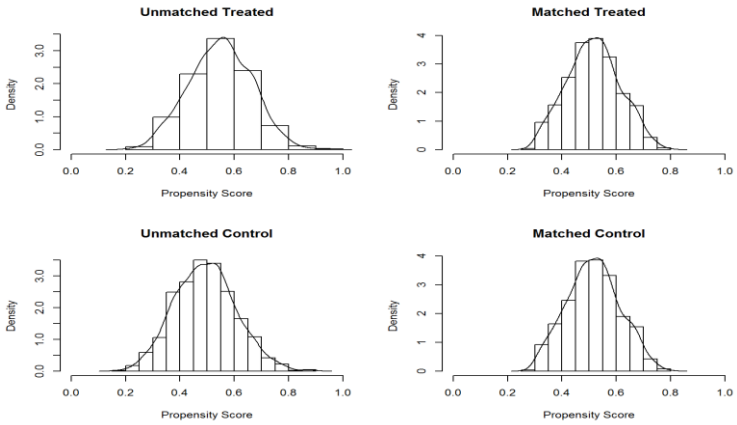
Relative multivariate imbalance L1 (Iacus, King, & Porro, 2010)

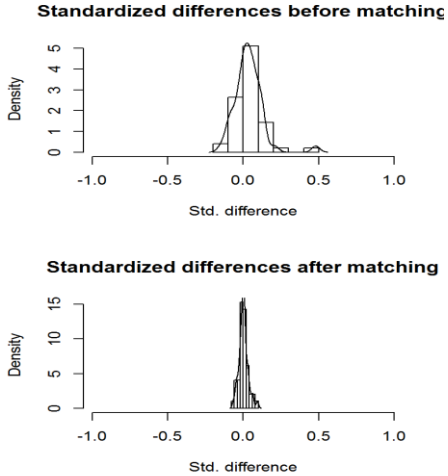
	Before matching	After matching
(all cases)	.999	.998

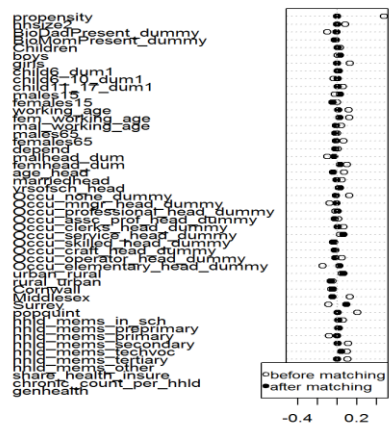
Summary of unbalanced covariates (|d| > .25)

No covariate exhibits a large imbalance (|d| > .25).









***PROPENSITY SCORE MATCHING NEAREST NEIGHBOR WITHOUT REPLACEMENT WITH CALIPER 0F **0.005** using all 48 VARIABLES**

Sample Sizes

	All		Matched		Unmatched		Discarded	
Subsamples	Control	Treated	Control	Treated	Control	Treated	Control	Treated
(all cases)	812	888	509	509	301	376	2	3

Overall balance test (Hansen & Bowers, 2010)

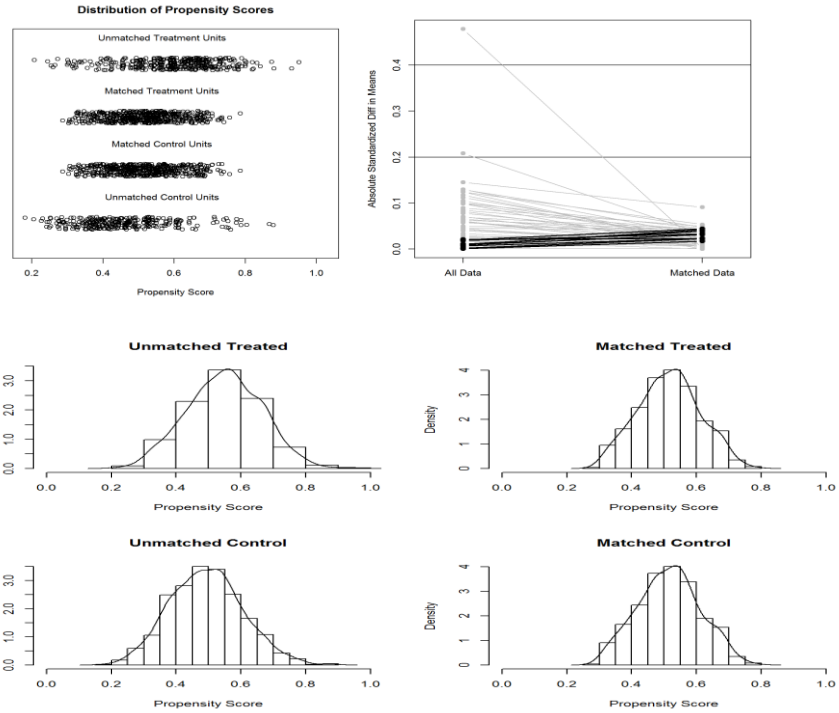
	chisquare	df	p.value
(all cases)	14.885	38.000	1.000

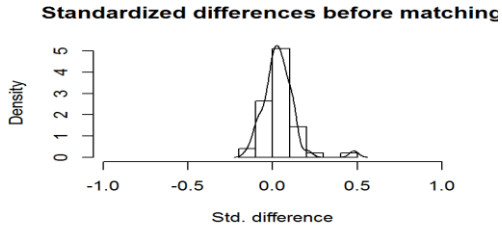
Relative multivariate imbalance L1 (Iacus, King, & Porro, 2010)

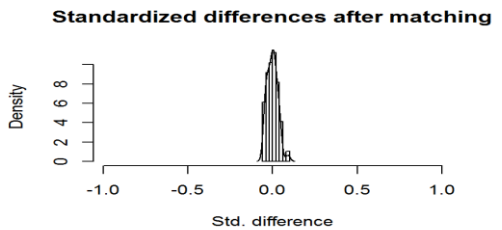
	Before matching	After matching
(all cases)	.999	.998

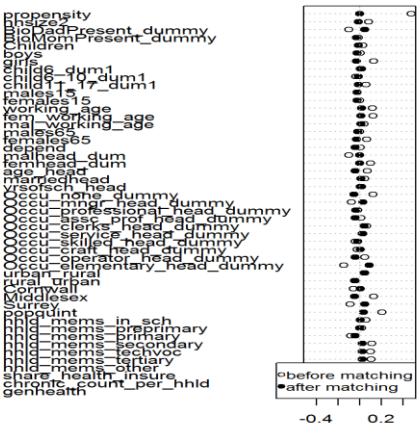
Summary of unbalanced covariates (|d| > .25)

No covariate exhibits a large imbalance (|d| > .25).









***PROPENSITY SCORE MATCHING NEAREST NEIGHBOR WITHOUT REPLACEMENT WITH CALIPER 0F **0.0025** using all 48 VARIABLES**

Sample Sizes

	All		Matched		Unmatched		Discarded	
Subsamples	Control	Treated	Control	Treated	Control	Treated	Control	Treated
(all cases)	812	888	421	421	389	464	2	3

Overall balance test (Hansen & Bowers, 2010)

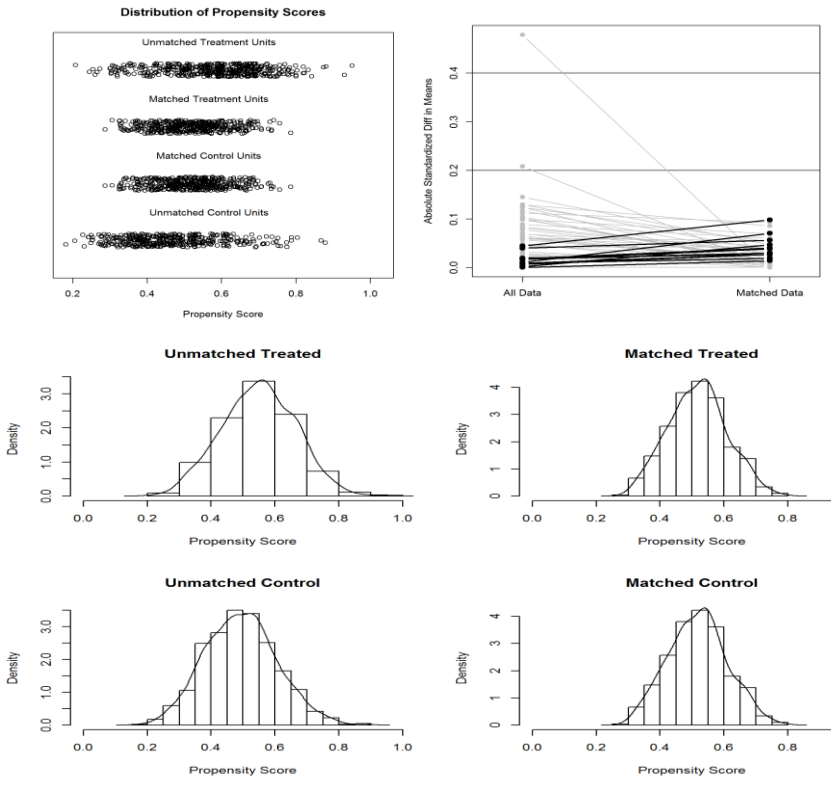
	chisquare	df	p.value
(all cases)	22.016	38.000	.982

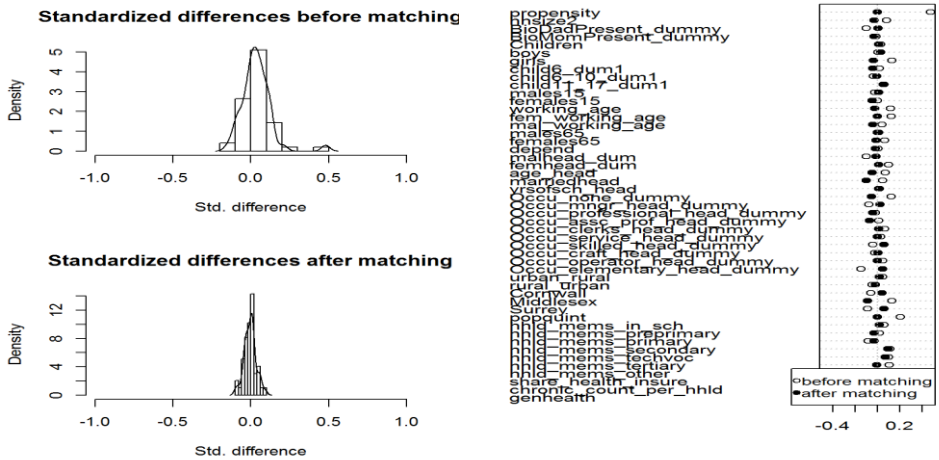
Relative multivariate imbalance L1 (Iacus, King, & Porro, 2010)

	Before matching	After matching
(all cases)	.999	.998

Summary of unbalanced covariates (|d| > .25)

No covariate exhibits a large imbalance (|d| > .25).





***PROPENSITY SCORE MATCHING NEAREST NEIGHBOR WITHOUT REPLACEMENT WITH CALIPER 0F 0.001 using all 48 VARIABLES**

Sample Sizes

	All		Matched		Unmatched		Discarded	
Subsamples	Control	Treated	Control	Treated	Control	Treated	Control	Treated
(all cases)	812	888	270	270	540	615	2	3

Overall balance test (Hansen & Bowers, 2010)

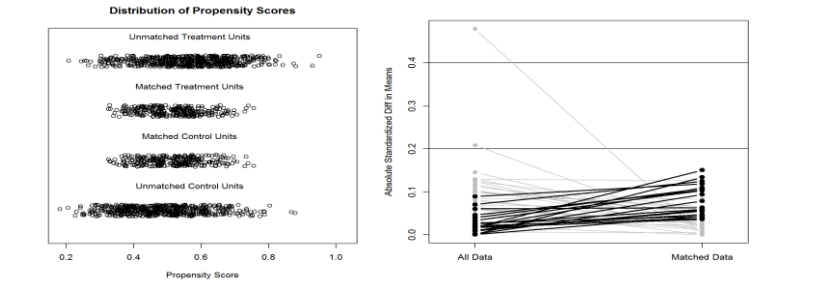
	chisquare	df	p.value
(all cases)	32.556	38.000	.719

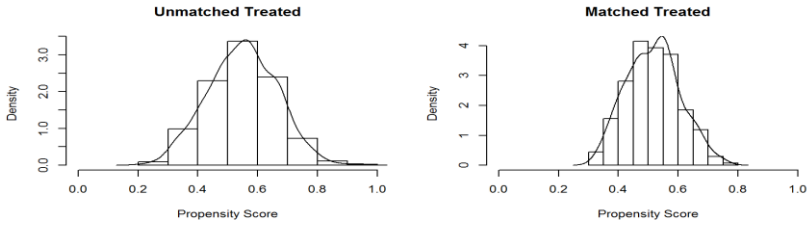
Relative multivariate imbalance L1 (Iacus, King, & Porro, 2010)

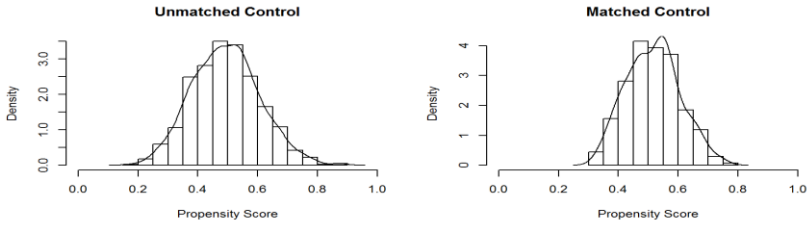
	Before matching	After matching
(all cases)	.999	.996

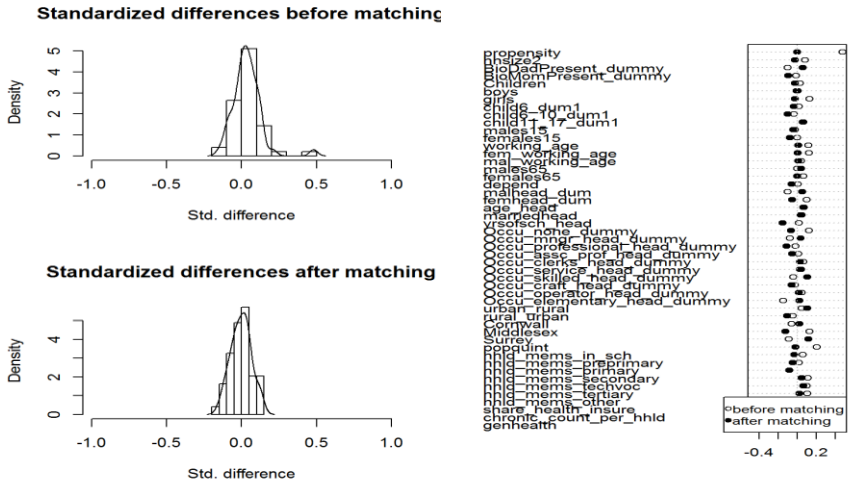
Summary of unbalanced covariates (|d| > .25)

No covariate exhibits a large imbalance (|d| > .25).









Appendix E: Analysis of results from matching with replacement

EDUCATION	RRH		NRRH		Difference
	Mean	Std. Dev	Mean	Std. Dev	
HOUSEHOLD CHARACTERISTICS					
Household Size >3	61 988.30	51 419.77	63 891.44	53 435.70	-1 903.14
All children in the household have their biological father present	70 723.69	67 209.47	71 274.66	63 655.57	-550.97
All children in the household have their biological mother present	63 196.69	60 077.96	60 692.98	54 230.96	2 503.70
Number of children > 0	69 045.89	54 923.93	70 067.96	57 157.69	-1 022.07
# of males under 18 years > 0	66 039.90	45 649.77	72 479.05	57 767.21	-6 439.15
# of females under 18 years >0	44 673.05	56 695.34	41 289.61	56 222.49	3 383.44
Number of children under 6 yrs > 0	48 454.12	44 154.55	43 105.31	33 576.50	5348.80*
Number of children 6-10 yrs > 0	70 689.27	51 898.45	69 675.85	47 938.52	1 013.42
Number of children 11-17 yrs > 0	80 201.39	47 981.91	86 928.43	60 735.05	-6 727.04
# of males below 15 years in household > 0	64 086.34	43 781.61	65 572.84	52 476.84	-1 486.50
# of females below 15 years in household	66 929.91	57 258.05	59 158.39	43 719.85	7771.51**
# of individuals age 15-64 > 0	37 957.00	54 342.33	35 441.69	53 897.65	2 515.30
# of females in working age group (15-64) > 0	49 461.54	57 290.23	45 736.98	56 608.69	3 724.56
# of males in working age group (15-64) > 0	35 297.77	51 958.29	34 115.10	54 596.93	1 182.67
# of males 65+ years in household > 0	16 839.65	37 537.10	20 422.04	49 274.03	-3 582.38
# of females 65+ years in household > 0	24 497.55	51 836.19	20 903.72	45 202.45	3 593.83
# of persons in dependent age group in household>0	48 434.23	55 172.38	46 804.59	53 720.65	1 629.64
EDUCATION	**RRH**		**NRRH**		**Difference**
	Mean	**Std. Dev**	**Mean**	**Std. Dev**	
HOUSEHOLD HEAD					
Male headed household	28 428.47	51 010.80	28 919.05	53 325.79	-490.59
Female headed household	41 121.46	54 016.59	36 078.69	50 440.34	5 042.78
Age of household head >64	17 074.97	38 134.37	17 255.90	41 017.09	-180.92
Married household head	45 877.58	61 144.89	40 795.75	62 277.56	5 081.83
Years of schooling of houshold head >12	45 418.65	63 596.44	34 466.58	57 059.31	10952.07***
Household head - no occupation	15 240.45	30 984.77	14 160.19	28 025.65	1 080.26
Household head Legislators, Senior Officials and Managers	61 874.18	74 947.57	47 671.02	71 628.18	14 203.17
Household head Professionals	70 152.44	83 649.93	53 866.70	87 912.24	16 285.74
Household head Technicians and Associate Professionals	60 588.89	74 300.09	45 246.51	65 345.78	15 342.38
Household head Clerks	47 948.56	53 545.21	24 116.15	39 560.39	23832.41***
Household head Service Workers and Shop and Market Sales Workers	39 987.26	48 375.10	34 869.66	44 772.02	5 117.60
Household head Skilled Agricultural and Fishery Workers	24 406.99	45 559.09	30 326.18	53 394.39	-5 919.19
Household head Craft and Related Trades Workers	30 133.45	49 352.79	32 554.50	44 254.80	-2 421.06
Household head Plant and Machine Operators and Assemblers	32 166.70	37 579.92	42 617.53	60 431.20	-10 450.83
Household head Elementary Occupations	28 953.84	40 875.03	33 937.16	43 759.53	-4 983.32
EDUCATION	**RRH**		**NRRH**		**Difference**
	Mean	**Std. Dev**	**Mean**	**Std. Dev**	
LOCATION					
Urban household	39 290.61	57 184.60	32 869.68	55 128.30	6420.92*
Rural household	28 805.44	47 116.36	31 165.77	49 234.39	-2 360.33
Cornwall County	31 679.02	47 039.32	29 533.89	44 253.74	2 145.13
Middlesex County	32 374.91	49 080.05	29 938.05	47 824.54	2 436.86
Surrey County	39 898.49	63 277.53	37 993.04	64 723.77	1 905.45
EDUCATION	**RRH**		**NRRH**		**Difference**
	Mean	**Std. Dev**	**Mean**	**Std. Dev**	
OTHER SOCIO-ECONOMIC					
Per Capita Population Quintile 1	29 424.85	30 316.97	32 592.21	36 773.05	-3 167.36
Per Capita Population Quintile 2	34 792.33	35 286.29	35 357.25	40 016.73	-564.92
Per Capita Population Quintile 3	38 735.61	49 144.44	37 777.40	40 887.71	958.21
Per Capita Population Quintile 4	42 818.14	60 307.70	35 745.69	62 343.56	7 072.44
Per Capita Population Quintile 5	27 545.85	60 575.04	24 466.82	58 186.30	3 079.03
Per Capita Population Quintile >2	34 693.07	58 011.76	31 291.79	56 267.21	3 401.27
# of household members in school	67 320.49	56 996.64	66 290.93	58 084.99	1 029.56
# of household members in pre-primary school	48 885.29	44 040.87	44 053.75	31 986.96	4 831.54
# of household members in primary school	72 837.91	50 802.11	73 124.55	54 744.48	-286.64
# of household members in secondary school	94 496.97	49 154.07	95 359.60	56 257.02	-862.63
# of household members in technical or vocational school	77 110.36	39 044.74	105 777.50	37 836.46	-28 667.14
# of household members in tertiary school	33 119.46	41 701.73	29 144.11	37 213.58	3 975.35
# of household members in other types of school	15 003.41	25 987.82	20 842.48	40 263.46	-5 839.07
Share of household members with health insurance	50 374.45	68 226.32	35 745.48	59 556.32	14628.97***
Number of chronic illnesses per houshold member	39 293.88	54 908.15	33 790.68	51 716.73	5 503.20
general health score of the household 1	28 121.22	50 634.76	38 137.68	64 360.35	-10 016.46
general health score of the household 2	31 973.31	52 731.06	23 179.25	40 647.21	8794.06***
general health score of the household 3	7 240.91	29 193.13	1 968.85	10 722.94	5272.06**
general health score of the household 4	0.00	0.00	0.00	0.00	0.00
general health score of the household 5	0.00	0.00	0.00	0.00	0.00
general health score of the household >2.5	6 833.53	25 981.42	6 167.13	18 253.81	666.39

HEALTH	RRH		NRRH		Difference
	Mean	Std. Dev	Mean	Std. Dev	
HOUSEHOLD CHARACTERISTICS					
Household Size >3	4,830.42	8,699.44	4,447.88	7,930.09	382.53
All children in the household have their biological father present	5,528.32	13,082.37	5,206.93	8,752.76	321.39
All children in the household have their biological mother present	6,320.20	15,607.38	4,824.71	9,184.31	1,495.49
Number of children > 0	5,945.19	13,422.17	5,098.93	8,662.76	846.26
# of males under 18 years > 0	5,306.59	11,496.37	5,368.32	9,232.46	-61.73
# of females under 18 years >0	8,883.45	17,026.99	8,181.83	17,396.16	701.62
Number of children under 6 yrs > 0	4,862.52	9,708.78	4,462.31	8,436.50	400.21
Number of children 6-10 yrs > 0	4,104.42	6,448.42	4,283.68	6,858.64	-179.26
Number of children 11-17 yrs > 0	6,423.57	15,211.17	4,446.54	7,622.02	1977.03**
# of males below 15 years in household > 0	5,128.07	10,115.25	5,273.33	9,522.82	-145.26
# of females below 15 years in household	5,372.77	12,891.40	3,911.68	5,788.20	1461.09**
# of individuals age 15-64 > 0	7,812.84	15,554.90	7,672.88	19,116.28	139.96
# of females in working age group (15-64) > 0	8,139.93	16,361.36	7,740.27	17,242.11	399.66
# of males in working age group (15-64) > 0	7,230.30	14,731.08	8,080.61	21,312.87	-850.32
# of males 65+ years in household > 0	13,145.84	26,712.67	11,155.88	15,261.02	1,989.97
# of females 65+ years in household > 0	11,178.64	16,322.48	9,349.68	15,269.92	1,828.96
# of persons in dependent age group in household >0	7,986.22	17,614.37	6,907.99	12,262.72	1,078.23
HEALTH	**RRH**		**NRRH**		**Difference**
	Mean	**Std. Dev**	**Mean**	**Std. Dev**	
HOUSEHOLD HEAD					
Male headed household	8,679.47	18,815.73	7,959.13	18,760.08	720.35
Female headed household	8,723.18	16,504.24	8,312.43	19,666.33	410.75
Age of household head >64	11,355.05	21,889.31	10,391.98	16,611.57	963.07
Married household head	11,683.10	23,785.84	9,571.34	15,376.78	2,111.76
Years of schooling of houshold head >12	9,922.06	18,902.72	8,816.60	19,043.66	1,105.46
Household head - no occupation	9,716.06	17,631.40	7,132.64	14,610.72	2583.42*
Household head Legislators, Senior Officials and Managers	10,535.89	16,535.26	22,262.46	47,255.29	-11,726.57
Household head Professionals	20,647.22	33,711.25	14,596.02	14,341.81	6051.20*
Household head Technicians and Associate Professionals	6,348.32	7,717.45	24,617.21	47,338.91	-18268.88**
Household head Clerks	10,901.33	17,425.45	6,372.68	7,823.40	4528.65**
Household head Service Workers and Shop and Market Sales Workers	5,761.04	9,879.86	3,614.96	6,011.11	2146.08***
Household head Skilled Agricultural and Fishery Workers	6,831.48	11,867.25	5,205.22	10,751.27	1,626.26
Household head Craft and Related Trades Workers	5,880.40	8,830.89	5,868.74	12,909.01	11.65
Household head Plant and Machine Operators and Assemblers	12,668.42	37,269.25	10,079.13	17,067.34	2,589.30
Household head Elementary Occupations	6,679.96	14,573.42	5,162.15	7,242.10	1,517.81
HEALTH	**RRH**		**NRRH**		**Difference**
	Mean	**Std. Dev**	**Mean**	**Std. Dev**	
LOCATION					
Urban household	10,301.20	20,388.76	10,317.02	24,695.00	-15.82
Rural household	7,020.37	14,430.54	5,976.97	11,076.44	1,043.40
Cornwall County	8,834.50	16,567.65	5,392.93	8,674.36	3441.58***
Middlesex County	9,310.41	19,382.03	7,783.23	19,604.67	1,527.18
Surrey County	7,436.72	15,772.25	11,260.59	24,401.25	-3823.86**
HEALTH	**RRH**		**NRRH**		**Difference**
	Mean	**Std. Dev**	**Mean**	**Std. Dev**	
OTHER SOCIO-ECONOMIC					
Per Capita Population Quintile 1	2,053.09	4,025.11	1,709.65	2,410.13	343.44
Per Capita Population Quintile 2	3,534.48	4,276.44	3,648.54	5,251.03	-114.06
Per Capita Population Quintile 3	5,546.00	9,999.07	6,055.42	12,910.38	-509.42
Per Capita Population Quintile 4	6,832.38	8,535.28	5,853.34	7,337.36	979.04
Per Capita Population Quintile 5	15,963.38	26,995.19	15,253.20	30,131.37	710.17
Per Capita Population Quintile >2	10,711.54	20,145.69	10,022.08	21,823.57	689.46
# of household members in school	6,568.98	14,386.62	5,479.07	9,181.68	1089.91*
# of household members in pre-primary school	4,996.01	9,814.54	4,435.79	8,499.75	560.21
# of household members in primary school	4,734.17	9,654.92	4,318.84	6,621.73	415.33**
# of household members in secondary school	6,410.53	14,949.43	4,416.33	7,231.20	1,994.20
# of household members in technical or vocational school	4,745.72	6,657.40	2,995.00	5,683.92	1,750.72
# of household members in tertiary school	11,871.78	23,023.09	9,333.86	12,597.09	2,537.91
# of household members in other types of school	7,821.71	11,361.74	6,390.69	8,737.13	1,431.02
Share of household members with health insurance	13,292.07	19,382.37	14,385.38	25,230.21	-1,093.31
Number of chronic illnesses per houshold member	10,601.15	15,498.37	10,270.39	16,303.85	330.76
general health score of the household 1	5,029.18	11,309.75	9,438.75	25,845.78	-4409.57*
general health score of the household 2	8,564.45	15,704.37	7,054.15	21,897.11	1,510.30
general health score of the household 3	14,588.93	29,313.01	13,361.35	16,480.59	1,227.59
general health score of the household 4	24,635.71	41,668.02	7,193.06	7,578.62	17442.66***
general health score of the household 5	29,000.00	0.00	100.00	0.00	28,900.00
general health score of the household >2.5	16,394.96	30,588.08	10,703.69	14,485.33	5691.27***

Appendix F: Actions taking to generate final dataset

Data manipulation of JSLC dataset

Action	**Count**
Datasets merged	14
New variables created including recodes	110
Variables in amalgamated dataset	242

Appendix G: SPSS Syntax for Variable and Dataset Generation, Recoding, Grouping, Matching and Statistical Tests

```
* Encoding: .
**SYNTAX FOR THESIS YEAR 1**

DATASET ACTIVATE DataSet1.
DO IF (K702 = 1).
RECODE HHOLDTYPE (SYSMIS=1) (ELSE=SYSMIS).
END IF.
EXECUTE.

DO IF (K704 = 1 or K706 = 1 or K708 = 1).
RECODE HHOLDTYPE (SYSMIS=2).
END IF.
EXECUTE.

RECODE HHOLDTYPE (SYSMIS=3).
EXECUTE.


**RECODE FOR REMITTANCE VS NON REMITTANCE RECEIVING **
DO IF (K702 = 1 or K704 = 1 or K706 = 1 or K708 = 1 ).
RECODE HHOLDTYPE2 (SYSMIS=1).
END IF.
EXECUTE.

RECODE HHOLDTYPE2 (SYSMIS=2).
EXECUTE.


**CHANGING HOUSEHOLD TYPE 2 - REMITTANCE RECEIVING VS NON-REMITTANCE RECEIVING REMITTANCE CATEGORIES FROM 1 ND 2 TO 1 AND 0 ***

DATASET ACTIVATE DataSet1.
RECODE HHOLDTYPE2 (1=1) (ELSE=0).
EXECUTE.


**FREQUENCIES FOR HHOLD**
FREQUENCIES VARIABLES=HHOLDTYPE
  /ORDER=ANALYSIS.

FREQUENCIES VARIABLES=HHOLDTYPE2
  /ORDER=ANALYSIS.


**recode remittance with $ dataset to remittance receiving left behind, other remittance receiving and non-remittance receiving**
DATASET ACTIVATE DataSet2.
DO IF (ITEM_CODEK = 702).
RECODE HHOLDTYPE (SYSMIS=1).
END IF.
EXECUTE.

DO IF (ITEM_CODEK = 704 or ITEM_CODEK = 706 or ITEM_CODEK = 708).
RECODE HHOLDTYPE (SYSMIS=2).
END IF.
EXECUTE.

RECODE HHOLDTYPE (SYSMIS=3).
EXECUTE.

**RECODE FOR  $ dataset REMITTANCE VS NON REMITTANCE RECEIVING **
DO IF (ITEM_CODEK = 702 or ITEM_CODEK = 704 or ITEM_CODEK = 706 or ITEM_CODEK = 708).
RECODE HHOLDTYPE2 (SYSMIS=1).
END IF.
EXECUTE.

RECODE HHOLDTYPE2 (SYSMIS=2).
EXECUTE.

**CALCULATING ANNUAL REMITTANCES**

RECODE HOWOFT_PERIOD ('1'=1) ('2'=7) ('3'=14) ('4'=30) ('5'=90) ('6'=180) ('7'=365)
   ('97'=SYSMIS) ('98'=SYSMIS) ('99'=SYSMIS) ('8'=1) ('9'=1) INTO HOWOFT_PERIOD_rec.
```

```
VARIABLE LABELS  HOWOFT_PERIOD_rec 'How often money received'.
EXECUTE.

*COMPUTING ANNUAL REMITTANCE*
COMPUTE ANN_OTR_INC=AMOUNT * (HOWOFT_TIME * (365/HOWOFT_PERIOD_rec)).
VARIABLE LABELS  ANN_OTR_INC 'Other Annual Income'.
EXECUTE.


**COMPUTING THE CONSUMPTION EXPENDITURE FOR THE HOUSEHOLDS BY USING THE INDIVIDUAL CONSUMPTIONS OF HOUSEHOLD MEMBERS**
**The individual consumptions for each items were turned into variables and the cell filled with the amount spent**

DATASET ACTIVATE DataSet7.
IF  (ITEM_CD = 3010) G3010=G4_r.
EXECUTE.

IF  (ITEM_CD = 3020) G3020=G4_r.
EXECUTE.

IF  (ITEM_CD = 3030) G3030=G4_r.
EXECUTE.

IF  (ITEM_CD = 3040) G3040=G4_r.
EXECUTE.

IF  (ITEM_CD = 3050) G3050=G4_r.
EXECUTE.

IF  (ITEM_CD = 3060) G3060=G4_r.
EXECUTE.

IF  (ITEM_CD = 3070) G3070=G4_r.
EXECUTE.

IF  (ITEM_CD = 3080) G3080=G4_r.
EXECUTE.

IF  (ITEM_CD = 3090) G3090=G4_r.
EXECUTE.

IF  (ITEM_CD = 3100) G3100=G4_r.
EXECUTE.

IF  (ITEM_CD = 3110) G3110=G4_r.
EXECUTE.

IF  (ITEM_CD = 3120) G3120=G4_r.
EXECUTE.

IF  (ITEM_CD = 3130) G3130=G4_r.
EXECUTE.

IF  (ITEM_CD = 3140) G3140=G4_r.
EXECUTE.

IF  (ITEM_CD = 3150) G3150=G4_r.
EXECUTE.

IF  (ITEM_CD = 3160) G3160=G4_r.
EXECUTE.

IF  (ITEM_CD = 3170) G3170=G4_r.
EXECUTE.

IF  (ITEM_CD = 3180) G3180=G4_r.
EXECUTE.

IF  (ITEM_CD = 3190) G3190=G4_r.
EXECUTE.

IF  (ITEM_CD = 3201) G3201=G4_r.
EXECUTE.
IF  (ITEM_CD = 3202) G3202=G4_r.
EXECUTE.
IF  (ITEM_CD = 3211) G3211=G4_r.
```

```
EXECUTE.
IF  (ITEM_CD = 3212) G3212=G4_r.
EXECUTE.
IF  (ITEM_CD = 3213) G3213=G4_r.
EXECUTE.
IF  (ITEM_CD = 3220) G3220=G4_r.
EXECUTE.
IF  (ITEM_CD = 3230) G3230=G4_r.
EXECUTE.
IF  (ITEM_CD = 3240) G3240=G4_r.
EXECUTE.
IF  (ITEM_CD = 3250) G3250=G4_r.
EXECUTE.
IF  (ITEM_CD = 3260) G3260=G4_r.
EXECUTE.
IF  (ITEM_CD = 3270) G3270=G4_r.
EXECUTE.
IF  (ITEM_CD = 3280) G3280=G4_r.
EXECUTE.
IF  (ITEM_CD = 3290) G3290=G4_r.
EXECUTE.
IF  (ITEM_CD = 3300) G3300=G4_r.
EXECUTE.
IF  (ITEM_CD = 3310) G3310=G4_r.
EXECUTE.
IF  (ITEM_CD = 3320) G3320=G4_r.
EXECUTE.
IF  (ITEM_CD = 3330) G3330=G4_r.
EXECUTE.
IF  (ITEM_CD = 3340) G3340=G4_r.
EXECUTE.
IF  (ITEM_CD = 3350) G3350=G4_r.
EXECUTE.
IF  (ITEM_CD = 3360) G3360=G4_r.
EXECUTE.
IF  (ITEM_CD = 3371) G3371=G4_r.
EXECUTE.
IF  (ITEM_CD = 3372) G3372=G4_r.
EXECUTE.
IF  (ITEM_CD = 3380) G3380=G4_r.
EXECUTE.
IF  (ITEM_CD = 3391) G3391=G4_r.
EXECUTE.
IF  (ITEM_CD = 3392) G3392=G4_r.
EXECUTE.
IF  (ITEM_CD = 3400) G3400=G4_r.
EXECUTE.
IF  (ITEM_CD = 3410) G3410=G4_r.
EXECUTE.
IF  (ITEM_CD = 3420) G3420=G4_r.
EXECUTE.
IF  (ITEM_CD = 3430) G3430=G4_r.
EXECUTE.
IF  (ITEM_CD = 3440) G3440=G4_r.
EXECUTE.
IF  (ITEM_CD = 3450) G3450=G4_r.
EXECUTE.
IF  (ITEM_CD = 3460) G3460=G4_r.
EXECUTE.
IF  (ITEM_CD = 3470) G3470=G4_r.
EXECUTE.
IF  (ITEM_CD = 3481) G3481=G4_r.
EXECUTE.
IF  (ITEM_CD = 3482) G3482=G4_r.
EXECUTE.
IF  (ITEM_CD = 3490) G3490=G4_r.
EXECUTE.
IF  (ITEM_CD = 3501) G3501=G4_r.
EXECUTE.
IF  (ITEM_CD = 3502) G3502=G4_r.
EXECUTE.
IF  (ITEM_CD = 3503) G3503=G4_r.
EXECUTE.
```

CALCULATING ANNUAL HOUSEHOLD EXPENDITURE ON HEALTH AND EDUCATION

```
DATASET ACTIVATE DataSet2.
COMPUTE Health=G3230+G3240+G3250.
VARIABLE LABELS  Health 'Annual Health Expenditure'.
EXECUTE.

COMPUTE Education=G3360+G3290.
VARIABLE LABELS  Education 'Annual Education Expenditure'.
EXECUTE.

FREQUENCIES VARIABLES=cons tot_food non_food Health Education
  /STATISTICS=MEAN SUM
  /ORDER=ANALYSIS.

FREQUENCIES VARIABLES=cons tot_food non_food Health Education utility housing
  /STATISTICS=MEAN SUM
  /ORDER=ANALYSIS.

FREQUENCIES VARIABLES=cons tot_food non_food Health Education utility housing per_cap1 per_cap2
  /STATISTICS=MEAN SUM
  /ORDER=ANALYSIS.

**RECODING HOUSEHOLD SIZE**
RECODE hhsize2_r (1=1) (2=2) (3=3) (4=4) (5 thru 10=5) (11 thru Highest=6).
EXECUTE.

RECODE hhsize2 (1=1) (2=2) (3=3) (4=4) (5 thru 10=5) (11 thru Highest=6) INTO hhsize2_r.
VARIABLE LABELS  hhsize2_r 'Household Size (grouped) - Members only'.
EXECUTE.

***AGGREGATE DESCROPTIVE TABLES***
* Custom Tables.
CTABLES
  /VLABELS VARIABLES=AREA hhsize2_r cons tot_food non_food utility housing Education Health
    HHOLDTYPE2
    DISPLAY=LABEL
  /TABLE AREA [C][COUNT F40.0] + hhsize2_r [C][COUNT F40.0] + cons [S][MEAN, SUM] + tot_food
    [S][MEAN, SUM] + non_food [S][MEAN, SUM] + utility [S][MEAN, SUM] + housing [S][MEAN, SUM] +
    Education [S][MEAN, SUM] + Health [S][MEAN, SUM] BY HHOLDTYPE2 [C]
  /CATEGORIES VARIABLES=AREA hhsize2_r HHOLDTYPE2 ORDER=A KEY=VALUE EMPTY=INCLUDE TOTAL=YES
    POSITION=AFTER
  /CRITERIA CILEVEL=95.


**GENERATING VARIABLE TO IDENTIFY CHILDREN IN HOUSEHOLD**
**use data set povline2015, Sect B_all.sav**
DATASET ACTIVATE DataSet1.
RECODE age (SYSMIS=SYSMIS) (Lowest thru 17=1) (18 thru Highest=2) INTO child.
VARIABLE LABELS  child 'Child or Not'.
EXECUTE.

RECODE child (1=1) (ELSE=SYSMIS) INTO child_dum1.
VARIABLE LABELS  child_dum1 'Child dummy'.
EXECUTE.

RECODE age (SYSMIS=SYSMIS) (6 thru Highest=2) (Lowest thru 5=1) INTO child10.
VARIABLE LABELS  child6 'Child under 6?'.
EXECUTE.

RECODE child6 (1=1) (ELSE=SYSMIS) INTO child6_dum1.
VARIABLE LABELS  child6_dum1 'Children under 6 dummy'.
EXECUTE.

DATASET ACTIVATE DataSet1.
RECODE age (SYSMIS=SYSMIS) (6 thru 10=1) (ELSE=2) INTO child6_10.
VARIABLE LABELS  child6_10 'Children 6-10 years old'.
EXECUTE.

RECODE child6_10 (1=1) (ELSE=SYSMIS) INTO child6_10_dum1.
VARIABLE LABELS  child6_10_dum1 'Children 6-10 years old dummy'.
EXECUTE.

DATASET ACTIVATE DataSet1.
RECODE age (SYSMIS=SYSMIS) (11 thru 17=1) (ELSE=2) INTO child11_17.
VARIABLE LABELS  child11_17 'Children 11-17 years old'.
EXECUTE.
```

```
RECODE child11_17 (1=1) (ELSE=SYSMIS) INTO child11_17_dum1.
VARIABLE LABELS  child11_17_dum1 'Children 11-17 years old dummy'.
EXECUTE.

RECODE age (SYSMIS=SYSMIS) (Lowest thru 14=1) (15 thru Highest=2) INTO child15.
VARIABLE LABELS  child15 'Children under 15 yrs'.
EXECUTE.

RECODE child15 (1=1) (ELSE=SYSMIS) INTO child15_dum1.
VARIABLE LABELS  child15_dum1 'Children under 15 yrs dummy'.
EXECUTE.

DATASET ACTIVATE DataSet1.
DO IF  (SEX = 1).
RECODE child15_dum1 (1=1) (ELSE=SYSMIS) INTO males15.
END IF.
VARIABLE LABELS  males15 'Males under 15 yrs old'.
EXECUTE.


DATASET ACTIVATE DataSet1.
DO IF  (SEX = 2).
RECODE child15_dum1 (1=1) (ELSE=SYSMIS) INTO females15.
END IF.
VARIABLE LABELS  females15 'Females under 15 yrs old'.
EXECUTE.

***GENERATING BOYS AND GIRLS***
DO IF  (SEX = 1).
RECODE child (1=1) (SYSMIS=0) (ELSE=0) INTO boys.
END IF.
VARIABLE LABELS  boys 'males under 18 years'.
EXECUTE.

RECODE boys (1=1) (SYSMIS=0) (ELSE=0).
EXECUTE.

DO IF  (SEX = 2).
RECODE child (SYSMIS=0) (2=1) (ELSE=0) INTO girls.
END IF.
VARIABLE LABELS  girls 'females under 18 years'.
EXECUTE.

RECODE girls (1=1) (SYSMIS=0) (ELSE=0).
EXECUTE.

**AGGREGATE FOR BOYS AND GIRLS***

DATASET DECLARE AggBoyGirl.
AGGREGATE OUTFILE='AggBoyGirl'
  /BREAK SERIAL
  /boys = sum(boys)
  /girls = sum(girls)

**GENERATING VAARIABLES RELATING TO THE SEX OF THE HOUSEHOLD HEAD**
**use data set povline2015, Sect B_all.sav**

IF  (RELAT = 1) sex_head=SEX.
VARIABLE LABELS  sex_head 'Sex of household head'.
EXECUTE.

RECODE sex_head (1=1) (ELSE=SYSMIS) INTO malhead_dum.
VARIABLE LABELS  malhead_dum 'male-headed household'.
EXECUTE.

RECODE sex_head (2=1) (ELSE=SYSMIS) INTO femhead_dum.
VARIABLE LABELS  femhead_dum 'female-headed household'.
EXECUTE.

**GENERATING VARIABLES RELATING TO THE AGE OF THE HOUSEHOLD HEAD**
**use data set povline2015, Sect B_all.sav**

IF  (RELAT = 1) age=head_age.
VARIABLE LABELS  age_head 'Age of household head'.
EXECUTE.
```

```
**GENERATING VARIABLES RELATING TO THE MARITAL STATUS OF THE HOUSEHOLD HEAD**
**use data set povline2015, Sect B_all.sav**

IF  (RELAT = 1) marstat_head=MARITAL_STAT.
VARIABLE LABELS  marstat_head 'Marital status of household head'.
EXECUTE.

***CREATING DUMMY FOR MARRIED HEAD OF HOUSHOLD***

DATASET ACTIVATE DataSet1.
RECODE marstat_head (1=1) (ELSE=0) INTO marriedhead.
VARIABLE LABELS  marriedhead 'Is the household head married?'.
EXECUTE.

**GENERATING VARIABLES RELATING TO THE UNION STATUS OF THE HOUSEHOLD HEAD**
**use data set povline2015, Sect B_all.sav**

IF  (RELAT = 1) unionstat_head=UNION_STAT.
VARIABLE LABELS  unionstat_head 'Union status of household head'.
EXECUTE.

***CREATING DUMMY FORHOUSEHOLD HEADS IN MARRITAL OR COHABITING UNION***

RECODE unionstat_head (1 thru 2=1) (ELSE=0) INTO unionhead_marr_cohab.
VARIABLE LABELS  unionhead_marr_cohab 'Is the houshold head in a marital or cohabiting union'.
EXECUTE.

**GENERATING VARIABLES RELATING TO THE POVERTY LINES OF THE HOUSEHOLD HEAD (ALL INDIVIDUALS IN EACH
HOUSHOL WOULD HAVE THE SAME SO THE POV LINE FOR EACH HOUSEHOLD HEAD WAS EXTRACTED**
**use data set povline2015, Sect B_all.sav**

IF  (RELAT = 1) povlinepercap_hhld=perline.
VARIABLE LABELS  povlinepercap_hhld 'Per capita poverty line for household'.
EXECUTE.

IF  (RELAT = 1) fpovlinepercap_hhld=fperline.
VARIABLE LABELS  fpovlinepercap_hhld 'Per capita food poverty line for household'.
EXECUTE.


**GENERATING VARIABLES RELATING TO THE EDUCATION OF THE HOUSEHOLD HEAD**
**use data set povline2015, Sect B_all.sav**

IF  (RELAT = 1) Edu_Attain_head=B24_SUBJECTS.
VARIABLE LABELS  Edu_Attain_head 'Highest educational attanment of  household head'.
EXECUTE.

IF  (RELAT = 1) yrsofsch_head=B23.
VARIABLE LABELS  yrsofsch_head 'Years of schooling of  household head'.
EXECUTE.

**GENERATING VARIABLE TO IDENTIFY WORKING AGE  PERSONS IN HOUSEHOLD**

**use data set povline2015, Sect B_all.sav**

DATASET ACTIVATE DataSet2.
RECODE age (SYSMIS=0) (15 thru 64=1) (ELSE=0) INTO working_age.
VARIABLE LABELS  working_age 'Individuals age 15-64 dummy'.
EXECUTE.

DO IF  (SEX = 1).
RECODE working_age (SYSMIS=0) (1=1) (ELSE=0) INTO mal_working_age.
END IF.
VARIABLE LABELS  mal_working_age 'Males in working age group (15-64)'.
EXECUTE.

DO IF  (SEX = 2).
RECODE working_age (SYSMIS=0) (1=1) (ELSE=0) INTO fem_working_age.
END IF.
VARIABLE LABELS  fem_working_age 'females in working age group (15-64)'.
EXECUTE.

RECODE mal_working_age fem_working_age (SYSMIS=0) (1=1).
EXECUTE.

**GENERATING AGGREGATE VARIABLES FOR WORKING AGE PERSONS*
```

```
**use data set povline2015, Sect B_all.sav**

DATASET DECLARE AggWrkingAge.
AGGREGATE OUTFILE='AggWrkingAge'
  /BREAK SERIAL
  /working_age = sum(working_age)
  /fem_working_age = sum(fem_working_age)
  /mal_working_age = sum(mal_working_age)


**GENERATING VARIABLE TO IDENTIFY ELDERLY PERSONS IN HOUSEHOLD**

**use data set povline2015, Sect B_all.sav**

DATASET ACTIVATE DataSet1.
RECODE age (SYSMIS=SYSMIS) (Lowest thru 64=2) (65 thru Highest=1) INTO elder.
VARIABLE LABELS  elder 'elderly age group (65+)'.
EXECUTE.

RECODE elder (1=1) (ELSE=SYSMIS) INTO elder_dum1.
VARIABLE LABELS  elder_dum1 'Elder dummy'.
EXECUTE.


DATASET ACTIVATE DataSet1.
DO IF  (SEX = 1).
RECODE elder_dum1 (1=1) (ELSE=SYSMIS) INTO males65.
END IF.
VARIABLE LABELS  males65 'Males 65 years and over'.
EXECUTE.


DATASET ACTIVATE DataSet1.
DO IF  (SEX = 2).
RECODE elder_dum1 (1=1) (ELSE=SYSMIS) INTO females65.
END IF.
VARIABLE LABELS  females65 'Females 65 years and over'.
EXECUTE.

DATASET ACTIVATE DataSet1.
RECODE age (Lowest thru 14=1) (65 thru Highest=1) (ELSE=0) INTO depend.
VARIABLE LABELS  depend 'Dependent houshold members (under 15 but over 64 yrs)'.
EXECUTE.

***GENERATING DUMMY FOR PRESENCE OF ELDERLY MALE OR ELDERLY FEMALE**
RECODE males65 females65 (0=0) (ELSE=1) INTO males65_present females65_present.
VARIABLE LABELS  males65_present 'Presence of males 65+ years' /females65_present 'Presence of '+
    'females 65+ years'.
EXECUTE.

**GENERATING AGGREGATE VARIABLES FOR THE TOTAL EDUCATIONAL EXPENDITURE ON VARIOUS ITEMS FOR EACH
HOUSEHOLD HEAD**
**use data set povline2015, Sect B_all.sav**

DATASET DECLARE AggEduExpenditure.
AGGREGATE OUTFILE='AggEduExpenditure'
  /BREAK SERIAL
  /exam_fees = sum(B29A)
  /tuition_incl_books = sum(B29B)
  /tuition_excl_books = sum(B29C)
  /auxillay_fees = sum(B29D)
  /extra_lessons = sum(B29E)
  /transport = sum(B29F)
  /lunch = sum(B29G)
  /uniform = sum(B29H)
  /books = sum(B29I)
  /other_supplies = sum(B29J)
  /boarding = sum(B29K)

DATASET ACTIVATE DataSet2.
COMPUTE EduExp=SUM(exam_fees,tuition_incl_books,tuition_excl_books,auxillay_fees,extra_lessons,
    transport,lunch,uniform,books,other_supplies,boarding).
VARIABLE LABELS  EduExp 'Expenditure on Education'.
EXECUTE.
```

```
**GENERATING AGGREGATE HOUSEHOLD VARIABLES FOR SELECTED VARIABLES IN THE POVLINE AND SECTION B DATASETS**
**use data set povline2015, Sect B_all.sav**

DATASET DECLARE AggPovline_and_SectB_all.
AGGREGATE OUTFILE='AggPovline_and_SectB_all'
  /BREAK SERIAL
 /Children = sum(child_dum1)
 /child6_dum1 = sum(child6_dum1)
 /child6_10_dum1 = sum(child6_10_dum1)
 /child11_17_dum1 = sum(child11_17_dum1)
 /sex_head = sum(sex_head)
 /malhead_dum = sum(malhead_dum)
 /femhead_dum = sum(femhead_dum)
 /age_head = sum(age_head)
 /marstat_head = sum(marstat_head)
 /unionstat_head = sum(unionstat_head)
 /povlinepercap_hhld = sum(povlinepercap_hhld)
 /fpovlinepercap_hhld = sum(fpovlinepercap_hhld)
 /Edu_Attain_head = sum(Edu_Attain_head)
 /yrsofsch_head = sum(yrsofsch_head)
 /child15_dum1 = sum(child15_dum1)
 /elder_dum1 = sum(elder_dum1)
 /males65 = sum(males65)
 /females65 = sum(females65)
 /males15 = sum(males15)
 /females15 = sum(females15)
 /depend = sum(depend)


 **GENERATING VARIABLES TO IDENTIFY HEALTH CHARACTERISTICS  IN HOUSEHOLD**

**use data set povline2015, SectA.sav**

DATASET ACTIVATE DataSet1.

RECODE health_Insurance (1=1) (ELSE=SYSMIS) INTO health_Insurance_dum1.
VARIABLE LABELS  health_Insurance_dum1 'Health Insurance dummy'.
EXECUTE.

** Calculating the share of the household with health insurance = number insured 7 hhld size)***

COMPUTE share_health_insure=health_Insurance / hhsize2.
VARIABLE LABELS  share_health_insure 'Share of household members with health insurance'.
EXECUTE.


**# of Chronic illnesses per houshold member**

DATASET ACTIVATE DataSet1.
RECODE A25_A ('1'=1) (ELSE=SYSMIS) INTO Asthma_dum1.
VARIABLE LABELS  Asthma_dum1 'Asthma dummy'.
EXECUTE.

DATASET ACTIVATE DataSet1.
RECODE A25_B ('1'=1) (ELSE=SYSMIS) INTO Diabetes_dum1.
VARIABLE LABELS  Diabetes_dum1 'Diabetes dummy'.
EXECUTE.

DATASET ACTIVATE DataSet1.
RECODE A25_C ('1'=1) (ELSE=SYSMIS) INTO Hypertension_dum1.
VARIABLE LABELS  Hypertension_dum1 'Hypertension dummy'.
EXECUTE.

DATASET ACTIVATE DataSet1.
RECODE A25_D ('1'=1) (ELSE=SYSMIS) INTO Arthritis_dum1.
VARIABLE LABELS  Arthritis_dum1 'Arthritis dummy'.
EXECUTE.

DATASET ACTIVATE DataSet1.
RECODE A25_E ('1'=1) (ELSE=SYSMIS) INTO MentalDisorder_dum1.
VARIABLE LABELS  MentalDisorder_dum1 'MentalDisorder dummy'.
EXECUTE.

DATASET ACTIVATE DataSet1.
RECODE A25_F ('1'=1) (ELSE=SYSMIS) INTO HeartDisease_dum1.
```

```
VARIABLE LABELS  HeartDisease_dum1 'HeartDisease dummy'.
EXECUTE.

DATASET ACTIVATE DataSet1.
RECODE A25_G ('1'=1) (ELSE=SYSMIS) INTO OtherChronic_dum1.
VARIABLE LABELS  OtherChronic_dum1 'OtherChronic dummy'.
EXECUTE.

COMPUTE chronic_count=SUM(Asthma_dum1,Diabetes_dum1,Hypertension_dum1,Arthritis_dum1,
    MentalDisorder_dum1,HeartDisease_dum1,OtherChronic_dum1).
VARIABLE LABELS  chronic_count 'Number of chronic illnesses'.
EXECUTE.

RECODE chronic_count (1=1) (2=2) (3=3) (4=4) (SYSMIS=0).
EXECUTE.

COMPUTE chronic_count_per_hhld=chronic_count / hhsize2.
VARIABLE LABELS  chronic_count_per_hhld 'Number of chronic illnesses per houshold member'.
EXECUTE.

IF  (chronic_count > 0) chronic_in_hhld=1.
VARIABLE LABELS  chronic_in_hhld 'Chronic illness present'.
EXECUTE.

  **GENERATING VARIABLES TO IDENTIFY ABSENT FATHER AND MOTHER FIGURE  IN HOUSEHOLD**

**use data set povline2015, SectA.sav**

DATASET ACTIVATE DataSet1.
DO IF  (age < 18).
RECODE FATH_FIG MOTH_FIG ('09'=1) (ELSE=0) INTO AbsentFather AbsentMother.
END IF.
VARIABLE LABELS  AbsentFather 'Father Figure Absent' /AbsentMother 'Mother Figure Absent'.
EXECUTE.
**1= father figure absent 0=father figure present**

DATASET ACTIVATE DataSet1.
DO IF  (age < 18).
RECODE FATH_FIG MOTH_FIG ('01'=0) (ELSE=1) INTO AbsentBioFather AbsentBioMother.
END IF.
VARIABLE LABELS  AbsentBioFather 'Father Biological Absent' /AbsentBioMother 'Mother Biological Absent'.
EXECUTE.
**1= father  absent 0=father  present**


***GENERATING DUMMY FOR WHETHER AT LEAST ONE CHILD IN THE HOUSEHOLD DOESNT HAVE A BIOLOGICAL
FATHER AND MOTHER PRESENT***

RECODE AbsentBioFather AbsentBioMother (0=0) (ELSE=1) INTO AbsBioFather_dummy AbsBioMother_dummy.
VARIABLE LABELS  AbsBioFather_dummy 'Is there atleast one child in the houshold with an absent '+
    'biological father?' /AbsBioMother_dummy 'Is there atleast one child in the houshold with an '+
    'absent biological mother?'.
EXECUTE.

**GENERATING THE NUMBER OF PRESENT BIOLOGICAL FATHERS AND MOTHERS IN HOUSEHOLD***

DATASET ACTIVATE Trimmed_Dataset_Year_1_Thesis.
COMPUTE PresentBioFather=Children - AbsentBioFather.
VARIABLE LABELS  PresentBioFather '# children with biological fathers present in household'.
EXECUTE.

IF  (Children > 0) RatioBioDadChild=(PresentBioFather / Children)*100.
VARIABLE LABELS  RatioBioDadChild '% Share of biological fathers to children present in household'.
EXECUTE.

RECODE RatioBioDadChild (100=1) (SYSMIS=0) (ELSE=0) INTO BioDadPresent_dummy.
VARIABLE LABELS  BioDadPresent_dummy 'All children in the household have their biological father '+
    'present'.
EXECUTE.


DATASET ACTIVATE Trimmed_Dataset_Year_1_Thesis.
COMPUTE PresentBioMother=Children - AbsentBioMother.
VARIABLE LABELS  PresentBioMother '# children with biological mothers present in household'.
EXECUTE.
```

```
IF  (Children > 0) RatioBioMomChild=(PresentBioMother / Children)*100.
VARIABLE LABELS  RatioBioMomChild '% Share of biological mothers to children present in household'.
EXECUTE.

RECODE RatioBioMomChild (100=1) (SYSMIS=0) (ELSE=0) INTO BioMomPresent_dummy.
VARIABLE LABELS  BioMomPresent_dummy 'All children in the household have their biological mother '+
    'present'.
EXECUTE.

IF  (BioDadPresent_dummy + BioMomPresent_dummy = 2) BothBioParents=1.
VARIABLE LABELS  BothBioParents 'Both parents of all children present in household are present'.
EXECUTE.

RECODE BothBioParents (1=1) (ELSE=0).
EXECUTE.

**GENERATING AGGREGATE HOUSEHOLD VARIABLES FOR SELECTED VARIABLES IN THE POVLINE AND SECTION A
DATASETS**
**use data set povline2015, SectA.sav**

     **this is the number of persons in the household with health insurance in the case of health insurance variable**

DATASET DECLARE AggPovline_and_SectA.
AGGREGATE OUTFILE='AggPovline_and_SectA'
  /BREAK SERIAL
  /genhealth = MEAN(A23)
  /health_Insurance = sum(health_Insurance_dum1)
  /Asthma = sum (Asthma_dum1)
  /Diabetes_dum1 = sum(Diabetes_dum1)
  /Hypertension_dum1 = sum(Hypertension_dum1)
  /Arthritis_dum1 = sum(Arthritis_dum1)
  /MentalDisorder_dum1 = sum(MentalDisorder_dum1)
  /HeartDisease_dum1 = sum(HeartDisease_dum1)
  /OtherChronic_dum1 = sum(OtherChronic_dum1)
  /chronic_count = sum(chronic_count)
  /chronic_in_hhld = sum(chronic_in_hhld)
  /AbsentFather = sum(AbsentFather)
  /AbsentMother = sum(AbsentMother)
  /AbsentBioFather = sum(AbsentBioFather)
  /AbsentBioMother = sum(AbsentBioMother)


  ** /Partner = SUM(Partner)
  /Child = SUM(Child)
  /Relative = SUM(Relative) **


***CALCULATING PER CAPITA DATA FOR SELECTED VARIABLES**

DATASET ACTIVATE DataSet1.
COMPUTE percap_t_noncon=t_noncon / hhsize2.
VARIABLE LABELS  percap_t_noncon 'Per Capita Annual Non-Consumption Expenditure'.
EXECUTE.


COMPUTE percap_non_food=non_food / hhsize2.
VARIABLE LABELS  percap_non_food 'Per Capita Annual Non-Food Expenditure'.
EXECUTE.

COMPUTE per_tot_food=tot_food / hhsize2.
VARIABLE LABELS  per_tot_food 'Per Capita Total Annual Food Expenditure'.
EXECUTE.

COMPUTE per_utility=utility / hhsize2.
VARIABLE LABELS  per_utility 'Per Capita Annual Utility Bill'.
EXECUTE.

COMPUTE per_housing=housing / hhsize2.
VARIABLE LABELS  per_housing 'Per Capita Annual Housing Expenditure'.
EXECUTE.

COMPUTE per_Health=Health / hhsize2.
VARIABLE LABELS  per_Health 'Per Capita Annual Health Expenditure'.
EXECUTE.

COMPUTE per_Education=Education / hhld_mems_in_sch.
```

```
VARIABLE LABELS  per_Education 'Per Capita Annual Education Expenditure'.
EXECUTE.

COMPUTE per_EducationAll=Education / hhsize2.
VARIABLE LABELS  per_EducationAll 'Per Capita Annual Education Expenditure based on all household member'.
EXECUTE.

COMPUTE per_exam_fees=exam_fees / hhld_mems_in_sch.
VARIABLE LABELS  per_exam_fees 'Per Capita exam_fees'.
EXECUTE.

COMPUTE per_tuition_excl_books=tuition_excl_books / hhld_mems_in_sch.
VARIABLE LABELS  per_tuition_excl_books 'Per Capita tuition_excl_books'.
EXECUTE.

COMPUTE per_tuition_incl_books=tuition_incl_books / hhld_mems_in_sch.
VARIABLE LABELS  per_tuition_incl_books 'Per Capita tuition_incl_books'.
EXECUTE.

COMPUTE per_auxillay_fees=auxillay_fees / hhld_mems_in_sch.
VARIABLE LABELS  per_auxillay_fees 'Per Capita auxillay_fees'.
EXECUTE.

COMPUTE per_extra_lessons=extra_lessons / hhld_mems_in_sch.
VARIABLE LABELS  per_extra_lessons 'Per Capita extra_lessons'.
EXECUTE.

COMPUTE per_transport=transport / hhld_mems_in_sch.
VARIABLE LABELS  per_transport 'Per Capita transport'.
EXECUTE.

COMPUTE per_lunch=lunch / hhld_mems_in_sch.
VARIABLE LABELS  per_lunch 'Per Capita lunch'.
EXECUTE.

COMPUTE per_uniform=uniform / hhld_mems_in_sch.
VARIABLE LABELS  per_uniform 'Per Capita uniform'.
EXECUTE.

COMPUTE per_books=books / hhld_mems_in_sch.
VARIABLE LABELS  per_books 'Per Capita books'.
EXECUTE.

COMPUTE per_other_supplies=other_supplies / hhld_mems_in_sch.
VARIABLE LABELS  per_other_supplies 'Per Capita other_supplies'.
EXECUTE.

COMPUTE per_boarding=boarding / hhld_mems_in_sch.
VARIABLE LABELS  per_boarding 'Per Capita boarding'.
EXECUTE.

COMPUTE per_EduExp=EduExp /  hhld_mems_in_sch.
VARIABLE LABELS  per_EduExp 'Per Capita EduExp'.
EXECUTE.


**RECODE TO GET WHETHER INDIVIDUAL IS GOING TO SOME FORM OF SCHOOL OR NOT***
**use data set Sect B_all.sav**

DATASET ACTIVATE DataSet3.
RECODE B1 (Lowest thru 18=1) (ELSE=0) INTO SCH_ATTENDANCE.
VARIABLE LABELS  SCH_ATTENDANCE 'Does the individual attend school'.
EXECUTE.


***RECODING TO CREADTE DUMMY FOR DIFFERENT TYPES OF SCHOOLING HOUSEHOLD MEMBERS WERE ATTENDING***
**use data set Sect B_all.sav**

RECODE B1 (Lowest thru 3=1) (ELSE=0) INTO SCH_Kinder.
VARIABLE LABELS  SCH_Kinder 'Individual attending below primary schooling'.
EXECUTE.

RECODE B1 (4 thru 7=1) (ELSE=0) INTO SCH_Primary.
VARIABLE LABELS  SCH_Primary 'Individual attending primary schooling'.
EXECUTE.
```

```
RECODE B1 (8 thru 9=1) (ELSE=0) INTO SCH_Secondary.
VARIABLE LABELS  SCH_Secondary 'Individual attending secondary schooling'.
EXECUTE.

RECODE B1 (10 thru 11=1) (ELSE=0) INTO SCH_TechVoc.
VARIABLE LABELS  SCH_TechVoc 'Individual attending technical or vocational schooling'.
EXECUTE.

RECODE B1 (12 thru 14=1) (ELSE=0) INTO SCH_Tertiary.
VARIABLE LABELS  SCH_Tertiary 'Individual attending tertiary schooling'.
EXECUTE.

RECODE B1 (15 thru 18=1) (ELSE=0) INTO SCH_Other.
VARIABLE LABELS  SCH_Other 'Individual attending other types of schooling'.
EXECUTE.


**GENERATING DUMMY VARIABLE FOR URBAN RURAL LOCATION AREA*****

DATASET ACTIVATE DataSet1.
RECODE AREA (SYSMIS=0) (1 thru 2=1) (ELSE=0) INTO urban_rural.
VARIABLE LABELS  urban_rural 'Location of household'.
EXECUTE.

RECODE urban_rural (0=1) (ELSE=0) INTO rural_urban.
VARIABLE LABELS  rural_urban 'Rural household'.
EXECUTE.

**GENERATING  VARIABLE FOR COUNTIES LOCATION AREA*****

DATASET ACTIVATE DataSet1.
RECODE PARISH (7 thru 11=1) (5 thru 6=2) (12 thru 14=2) (1 thru 4=3) INTO county.
VARIABLE LABELS  county 'County'.
EXECUTE.


**GENERATING  DUMMY VARIABLE FOR COUNTIES LOCATION AREA*****

RECODE county (1=1) (ELSE=0) INTO Cornwall.
VARIABLE LABELS  Cornwall 'Cornwall County'.
EXECUTE.

RECODE county (2=1) (ELSE=0) INTO Middlesex.
VARIABLE LABELS  Middlesex 'Middlesex County'.
EXECUTE.

RECODE county (3=1) (ELSE=0) INTO Surrey.
VARIABLE LABELS  Surrey 'Surrey County'.
EXECUTE.

**GENERATING AGGREGATE VARIABLES FOR THE TOTAL PERSONS IN HOUSEHOLD WHO ARE IN SOME FORM OF
SCHOOL**
**use data set Sect B_all.sav**

DATASET DECLARE AggAttendSchool.
AGGREGATE OUTFILE='AggAttendSchool'
  /BREAK SERIAL
  /hhld_mems_in_sch = sum(SCH_ATTENDANCE)
  /hhld_mems_preprimary = sum(SCH_Kinder)
  /hhld_mems_primary = sum(SCH_Primary)
  /hhld_mems_secondary = sum(SCH_Secondary)
  /hhld_mems_techvoc = sum(SCH_TechVoc)
  /hhld_mems_tertiary = sum(SCH_Tertiary)
  /hhld_mems_other = sum(SCH_Other)


***GENERATING DUMMY VARIABLES FOR THE DIFERRENT OCCUPATIONS***

DATASET ACTIVATE DataSet1.
RECODE P_EarnerOCCUPATION_R (0=1) (ELSE=0) INTO Occu_none_dummy.
VARIABLE LABELS  Occu_none_dummy 'Household head - no occupation'.
EXECUTE.
```

```
RECODE P_EarnerOCCUPATION_R (1=1) (ELSE=0) INTO Occu_mngr_head_dummy.
VARIABLE LABELS  Occu_mngr_head_dummy 'Household head  Legislators, Senior Officials and Managers '.
EXECUTE.

RECODE P_EarnerOCCUPATION_R (2=1) (ELSE=0) INTO Occu_professional_head_dummy.
VARIABLE LABELS  Occu_professional_head_dummy 'Household head Professionals '.
EXECUTE.

RECODE P_EarnerOCCUPATION_R (3=1) (ELSE=0) INTO Occu_assc_prof_head_dummy.
VARIABLE LABELS  Occu_assc_prof_head_dummy 'Household head Technicians and Associate Professionals  '.
EXECUTE.

RECODE P_EarnerOCCUPATION_R (4=1) (ELSE=0) INTO Occu_clerks_head_dummy.
VARIABLE LABELS  Occu_clerks_head_dummy 'Household head Clerks  '.
EXECUTE.

RECODE P_EarnerOCCUPATION_R (5=1) (ELSE=0) INTO Occu_service_head_dummy.
VARIABLE LABELS  Occu_service_head_dummy 'Household head Service Workers and Shop and Market '+
    'Sales Workers   '.
EXECUTE.

RECODE P_EarnerOCCUPATION_R (6=1) (ELSE=0) INTO Occu_skilled_head_dummy.
VARIABLE LABELS  Occu_skilled_head_dummy 'Household head Skilled Agricultural and Fishery Workers '.
EXECUTE.

RECODE P_EarnerOCCUPATION_R (7=1) (ELSE=0) INTO Occu_craft_head_dummy.
VARIABLE LABELS  Occu_craft_head_dummy 'Household head Craft and Related Trades Workers '.
EXECUTE.

RECODE P_EarnerOCCUPATION_R (8=1) (ELSE=0) INTO Occu_operator_head_dummy.
VARIABLE LABELS  Occu_operator_head_dummy 'Household head Plant and Machine Operators and Assemblers'.
EXECUTE.

RECODE P_EarnerOCCUPATION_R (9=1) (ELSE=0) INTO Occu_elementary_head_dummy.
VARIABLE LABELS  Occu_elementary_head_dummy 'Household head Elementary Occupations'.
EXECUTE.


***GROUPING 9 OCCUPATIONAL CATEGORIES TO 5
Managers & Professionals
Clersks & Service Workers
Skilled, Craftsmen & Machine Operators
Elementary Occupations**

DATASET ACTIVATE DataSet1.
RECODE P_EarnerOCCUPATION_R (0=0) (9=5) (8=4)(1 thru 3=1) (4 thru 5=2) (6 thru 7=3) INTO
    Occupation_R_Grouped.
VARIABLE LABELS  Occupation_R_Grouped 'Occupation of Household Head (Grouped)'.
EXECUTE.

**DUMMY VARIABLES FOR ABOVE OCCUPATION GROUPS***

IF  (Occupation_R_Grouped=0) GrpOccuHead_NoOccup=1.
VARIABLE LABELS  GrpOccuHead_NoOccup 'Head with No Occupation'.
EXECUTE.

IF  (Occupation_R_Grouped=1) GrpOccuHead_MngrProf=1.
VARIABLE LABELS  GrpOccuHead_MngrProf 'Heads who are Managers, Professionals or Associates'.
EXECUTE.

IF  (Occupation_R_Grouped=2) GrpOccuHead_ClerksSvceWrk=1.
VARIABLE LABELS  GrpOccuHead_ClerksSvceWrk 'Heads who are Clerks or Services Workers'.
EXECUTE.

IF  (Occupation_R_Grouped=3) GrpOccuHead_SkillCrafts=1.
VARIABLE LABELS  GrpOccuHead_SkillCrafts 'Heads who are Skilled and Craftsmen'.
EXECUTE.

IF  (Occupation_R_Grouped=4) GrpOccuHead_Op=1.
VARIABLE LABELS  GrpOccuHead_Op 'Heads who are machine operators'.
EXECUTE.

IF  (Occupation_R_Grouped=5) GrpOccuHead_Elem=1.
VARIABLE LABELS  GrpOccuHead_Elem 'Heads who have elementary occupations'.
EXECUTE.
```

```
RECODE GrpOccuHead_NoOccup GrpOccuHead_MngrProf GrpOccuHead_ClerksSvceWrk GrpOccuHead_SkillCrafts  GrpOccuHead_Op
GrpOccuHead_Elem (1=1) (SYSMIS=0) (ELSE=0).
EXECUTE.


**PROPENSITY SCORE MATCHING**

PSMATCHING3
  /VARS
     TREAT = HHOLDTYPE2
     COVS = PARISH CONST AREA popquint hhsize2_r CARE_GIVER P_EarnerOCCUPATION_R
    P_EarnerINDUSTRY_R Children child6_dum1 child6_10_dum1 child11_17_dum1 malhead_dum femhead_dum
    age_head marstat_head unionstat_head povlinepercap_hhld fpovlinepercap_hhld yrsofsch_head
    child15_dum1 elder_dum1 males65 females65 males15 females15 depend genhealth health_Insurance
    chronic_in_hhld AbsentFather AbsentMother AbsentBioFather AbsentBioMother
  /MATCHIT
     MATCH=NEAREST
     EST =LOGIT
     DISCARD = NONE
     MORDER = LARGEST
     RATIO = 1
     CALIPER = .2
     INTERACTION
  /PLOT HISTPLOT JITTERPLOT HISTBAL DOTPLOT INDBAL RESOLUTION = 96
  /OUTPUT PS PAIRED MATCHED_CASES_ONLY.



**PROPENSITY SCORE MATCHING WITH CALIPER 0.1**

PSMATCHING3
  /VARS
     TREAT = HHOLDTYPE2
     COVS = PARISH CONST AREA popquint hhsize2_r CARE_GIVER P_EarnerOCCUPATION_R
    P_EarnerINDUSTRY_R Children child6_dum1 child6_10_dum1 child11_17_dum1 malhead_dum femhead_dum
    age_head marstat_head unionstat_head povlinepercap_hhld fpovlinepercap_hhld yrsofsch_head
    child15_dum1 elder_dum1 males65 females65 males15 females15 depend genhealth health_Insurance
    chronic_in_hhld AbsentFather AbsentMother AbsentBioFather AbsentBioMother
  /MATCHIT
     MATCH=NEAREST
     EST =LOGIT
     DISCARD = NONE
     MORDER = LARGEST
     RATIO = 1
     CALIPER = .1
     INTERACTION
  /PLOT HISTPLOT JITTERPLOT HISTBAL DOTPLOT INDBAL RESOLUTION = 96
  /OUTPUT PS PAIRED MATCHED_CASES_ONLY.


**PROPENSITY SCORE MATCHING WITH CALIPER 0.05**

PSMATCHING3
  /VARS
     TREAT = HHOLDTYPE2
     COVS = PARISH CONST AREA popquint hhsize2_r CARE_GIVER P_EarnerOCCUPATION_R
    P_EarnerINDUSTRY_R Children child6_dum1 child6_10_dum1 child11_17_dum1 malhead_dum femhead_dum
    age_head marstat_head unionstat_head povlinepercap_hhld fpovlinepercap_hhld yrsofsch_head
    child15_dum1 elder_dum1 males65 females65 males15 females15 depend genhealth health_Insurance
    chronic_in_hhld AbsentFather AbsentMother AbsentBioFather AbsentBioMother
  /MATCHIT
     MATCH=NEAREST
     EST =LOGIT
     DISCARD = NONE
     MORDER = LARGEST
     RATIO = 1
     CALIPER = .05
     INTERACTION
  /PLOT HISTPLOT JITTERPLOT HISTBAL DOTPLOT INDBAL RESOLUTION = 96
  /OUTPUT PS PAIRED MATCHED_CASES_ONLY.


**PROPENSITY SCORE MATCHING WITH CALIPER 0.01**
PSMATCHING3
  /VARS
     TREAT = HHOLDTYPE2
```

```
    COVS = PARISH CONST AREA popquint hhsize2_r CARE_GIVER P_EarnerOCCUPATION_R
   P_EarnerINDUSTRY_R Children child6_dum1 child6_10_dum1 child11_17_dum1 malhead_dum femhead_dum
   age_head marstat_head unionstat_head povlinepercap_hhld fpovlinepercap_hhld yrsofsch_head
   child15_dum1 elder_dum1 males65 females65 males15 females15 depend genhealth health_Insurance
   chronic_in_hhld AbsentFather AbsentMother AbsentBioFather AbsentBioMother
  /MATCHIT
    MATCH=NEAREST
    EST =LOGIT
    DISCARD = NONE
    MORDER = LARGEST
    RATIO = 1
    CALIPER = .01
    INTERACTION
  /PLOT HISTPLOT JITTERPLOT HISTBAL DOTPLOT INDBAL RESOLUTION = 96
  /OUTPUT PS PAIRED MATCHED_CASES_ONLY.


**PROPENSITY SCORE MATCHING WITH CALIPER 0.0001**
PSMATCHING3
  /VARS
    TREAT = HHOLDTYPE2
    COVS = PARISH CONST AREA popquint hhsize2_r CARE_GIVER P_EarnerOCCUPATION_R
   P_EarnerINDUSTRY_R Children child6_dum1 child6_10_dum1 child11_17_dum1 malhead_dum femhead_dum
   age_head marstat_head unionstat_head povlinepercap_hhld fpovlinepercap_hhld yrsofsch_head
   child15_dum1 elder_dum1 males65 females65 males15 females15 depend genhealth health_Insurance
   chronic_in_hhld AbsentFather AbsentMother AbsentBioFather AbsentBioMother
  /MATCHIT
    MATCH=NEAREST
    EST =LOGIT
    DISCARD = NONE
    MORDER = LARGEST
    RATIO = 1
    CALIPER = .0001
    INTERACTION
  /PLOT HISTPLOT JITTERPLOT HISTBAL DOTPLOT INDBAL RESOLUTION = 96
  /OUTPUT PS PAIRED MATCHED_CASES_ONLY.


**PROPENSITY SCORE MATCHING WITH CALIPER 0.001**
PSMATCHING3
  /VARS
    TREAT = HHOLDTYPE2
    COVS = PARISH CONST AREA popquint hhsize2_r CARE_GIVER P_EarnerOCCUPATION_R
   P_EarnerINDUSTRY_R Children child6_dum1 child6_10_dum1 child11_17_dum1 malhead_dum femhead_dum
   age_head marstat_head unionstat_head povlinepercap_hhld fpovlinepercap_hhld yrsofsch_head
   child15_dum1 elder_dum1 males65 females65 males15 females15 depend genhealth health_Insurance
   chronic_in_hhld AbsentFather AbsentMother AbsentBioFather AbsentBioMother
  /MATCHIT
    MATCH=NEAREST
    EST =LOGIT
    DISCARD = NONE
    MORDER = LARGEST
    RATIO = 1
    CALIPER = .001
    INTERACTION
  /PLOT HISTPLOT JITTERPLOT HISTBAL DOTPLOT INDBAL RESOLUTION = 96
  /OUTPUT PS PAIRED MATCHED_CASES_ONLY.



**PROPENSITY SCORE MATCHING WITH CALIPER 0.001 (EXCLUDING INDUSTRY OF PRINCIPAL EARNER**
PSMATCHING3
  /VARS
    TREAT = HHOLDTYPE2
    COVS = PARISH CONST AREA popquint hhsize2_r CARE_GIVER P_EarnerOCCUPATION_R
   Children child6_dum1 child6_10_dum1 child11_17_dum1 malhead_dum femhead_dum
   age_head marstat_head unionstat_head povlinepercap_hhld fpovlinepercap_hhld yrsofsch_head
   child15_dum1 elder_dum1 males65 females65 males15 females15 depend genhealth health_Insurance
   chronic_in_hhld AbsentFather AbsentMother AbsentBioFather AbsentBioMother
  /MATCHIT
    MATCH=NEAREST
    EST =LOGIT
    DISCARD = NONE
    MORDER = LARGEST
    RATIO = 1
    CALIPER = .001
```

```
    INTERACTION
  /PLOT HISTPLOT JITTERPLOT HISTBAL DOTPLOT INDBAL RESOLUTION = 96
  /OUTPUT PS PAIRED MATCHED_CASES_ONLY.


***PSM NEAREST NEIGHBOUR WITHOUT REPLACEMENT ON CALIPER 0.0001***
PSMATCHING3
  /VARS
    TREAT = HHOLDTYPE2
    COVS = yrsofsch_head marriedhead femhead_dum BioDadPresent_dummy BioMomPresent_dummy boys
   hhsize2 popquint urban_rural hhld_mems_primary hhld_mems_secondary hhld_mems_tertiary
  /MATCHIT
    MATCH=NEAREST
    EST =LOGIT
    DISCARD = NONE
    MORDER = LARGEST
    RATIO = 1
    CALIPER = .0001
  /PLOT HISTPLOT JITTERPLOT HISTBAL DOTPLOT INDBAL RESOLUTION = 96
  /OUTPUT PS PAIRED MATCHED_CASES_ONLY.



***PSM NEAREST NEIGHBOUR WITHOUT REPLACEMENT ON CALIPER 0.001***
PSMATCHING3
  /VARS
    TREAT = HHOLDTYPE2
    COVS = yrsofsch_head marriedhead femhead_dum BioDadPresent_dummy BioMomPresent_dummy boys
   hhsize2 popquint urban_rural hhld_mems_primary hhld_mems_secondary hhld_mems_tertiary
  /MATCHIT
    MATCH=NEAREST
    EST =LOGIT
    DISCARD = NONE
    MORDER = LARGEST
    RATIO = 1
    CALIPER = .001
  /PLOT HISTPLOT JITTERPLOT HISTBAL DOTPLOT INDBAL RESOLUTION = 96
  /OUTPUT PS PAIRED MATCHED_CASES_ONLY.


DATASET ACTIVATE ps_20190612192812.
ONEWAY yrsofsch_head marriedhead femhead_dum BioDadPresent_dummy BioMomPresent_dummy boys hhsize2
    popquint urban_rural hhld_mems_primary hhld_mems_secondary hhld_mems_tertiary BY HHOLDTYPE2
  /STATISTICS DESCRIPTIVES EFFECTS HOMOGENEITY BROWNFORSYTHE WELCH
  /PLOT MEANS
  /MISSING ANALYSIS.


****REGRESSION*****

PLUM per_EducationAll BY HHOLDTYPE2 WITH yrsofsch_head marriedhead femhead_dum BioDadPresent_dummy
    BioMomPresent_dummy boys girls Children hhsize2 popquint urban_rural hhld_mems_primary
    hhld_mems_secondary hhld_mems_tertiary
  /CRITERIA=CIN(95) DELTA(0) LCONVERGE(0) MXITER(100) MXSTEP(5) PCONVERGE(1.0E-6) SINGULAR(1.0E-8)
  /LINK=LOGIT
  /PRINT=FIT PARAMETER SUMMARY.


***TTEST ON COVARIATES**

T-TEST GROUPS=HHOLDTYPE2(1 0)
  /MISSING=ANALYSIS
  /VARIABLES=hhsize2 BioDadPresent_dummy BioMomPresent_dummy Children boys girls child6_dum1
    child6_10_dum1 child11_17_dum1 males15 females15 working_age mal_working_age fem_working_age
    males65 females65 depend malhead_dum femhead_dum age_head marriedhead yrsofsch_head Occu_none_dummy
    Occu_mngr_head_dummy Occu_professional_head_dummy Occu_assc_prof_head_dummy Occu_clerks_head_dummy
    Occu_service_head_dummy Occu_skilled_head_dummy Occu_craft_head_dummy Occu_operator_head_dummy
    Occu_elementary_head_dummy urban_rural rural_urban Cornwall Middlesex Surrey popquint
    hhld_mems_in_sch hhld_mems_preprimary hhld_mems_primary hhld_mems_secondary hhld_mems_techvoc
    hhld_mems_tertiary hhld_mems_other share_health_insure chronic_count_per_hhld genhealth
  /CRITERIA=CI(.99).


T-TEST GROUPS=HHOLDTYPE2(1 0)
  /MISSING=ANALYSIS
  /VARIABLES=hhsize2 BioDadPresent_dummy BioMomPresent_dummy Children boys girls child6_dum1
```

```
  child6_10_dum1 child11_17_dum1 males15 females15 working_age mal_working_age fem_working_age
  males65 females65 depend malhead_dum femhead_dum age_head marriedhead yrsofsch_head Occu_none_dummy
  Occu_mngr_head_dummy Occu_professional_head_dummy Occu_assc_prof_head_dummy Occu_clerks_head_dummy
  Occu_service_head_dummy Occu_skilled_head_dummy Occu_craft_head_dummy Occu_operator_head_dummy
  Occu_elementary_head_dummy urban_rural rural_urban Cornwall Middlesex Surrey popquint
  hhld_mems_in_sch hhld_mems_preprimary hhld_mems_primary hhld_mems_secondary hhld_mems_techvoc
  hhld_mems_tertiary hhld_mems_other share_health_insure chronic_count_per_hhld genhealth
 /CRITERIA=CI(.95).

T-TEST GROUPS=HHOLDTYPE2(1 0)
 /MISSING=ANALYSIS
 /VARIABLES=hhsize2 BioDadPresent_dummy BioMomPresent_dummy Children boys girls child6_dum1
  child6_10_dum1 child11_17_dum1 males15 females15 working_age mal_working_age fem_working_age
  males65 females65 depend malhead_dum femhead_dum age_head marriedhead yrsofsch_head Occu_none_dummy
  Occu_mngr_head_dummy Occu_professional_head_dummy Occu_assc_prof_head_dummy Occu_clerks_head_dummy
  Occu_service_head_dummy Occu_skilled_head_dummy Occu_craft_head_dummy Occu_operator_head_dummy
  Occu_elementary_head_dummy urban_rural rural_urban Cornwall Middlesex Surrey popquint
  hhld_mems_in_sch hhld_mems_preprimary hhld_mems_primary hhld_mems_secondary hhld_mems_techvoc
  hhld_mems_tertiary hhld_mems_other share_health_insure chronic_count_per_hhld genhealth
 /CRITERIA=CI(.90).


***PROPENSITY SCORE MATCHING NEAREST NEIGHBOR WITHOUT REPLACEMENT WITH CALIPER 0F 0.0001 using all 48
VARIABLES**

DATASET ACTIVATE DataSet1.
PSMATCHING3
  /VARS
    TREAT = HHOLDTYPE2
    COVS = hhsize2 BioDadPresent_dummy BioMomPresent_dummy Children boys girls child6_dum1
  child6_10_dum1 child11_17_dum1 males15 females15 working_age fem_working_age mal_working_age
  males65 females65 depend malhead_dum femhead_dum age_head marriedhead yrsofsch_head Occu_none_dummy
  Occu_mngr_head_dummy Occu_professional_head_dummy Occu_assc_prof_head_dummy Occu_clerks_head_dummy
  Occu_service_head_dummy Occu_skilled_head_dummy Occu_craft_head_dummy Occu_operator_head_dummy
  Occu_elementary_head_dummy urban_rural rural_urban Cornwall Middlesex Surrey popquint
  hhld_mems_in_sch hhld_mems_preprimary hhld_mems_primary hhld_mems_secondary hhld_mems_techvoc
  hhld_mems_tertiary hhld_mems_other share_health_insure chronic_count_per_hhld genhealth
  /MATCHIT
    MATCH=NEAREST
    EST =LOGIT
    DISCARD = BOTH
    MORDER = LARGEST
    RATIO = 1
    CALIPER = .0001
  /PLOT HISTPLOT JITTERPLOT HISTBAL DOTPLOT INDBAL RESOLUTION = 200
  /OUTPUT PAIRED MATCHED_CASES_ONLY.


***PROPENSITY SCORE MATCHING NEAREST NEIGHBOR WITHOUT REPLACEMENT WITH CALIPER 0F 0.0075 using all 48
VARIABLES**

DATASET ACTIVATE DataSet1.
PSMATCHING3
  /VARS
    TREAT = HHOLDTYPE2
    COVS = hhsize2 BioDadPresent_dummy BioMomPresent_dummy Children boys girls child6_dum1
  child6_10_dum1 child11_17_dum1 males15 females15 working_age fem_working_age mal_working_age
  males65 females65 depend malhead_dum femhead_dum age_head marriedhead yrsofsch_head Occu_none_dummy
  Occu_mngr_head_dummy Occu_professional_head_dummy Occu_assc_prof_head_dummy Occu_clerks_head_dummy
  Occu_service_head_dummy Occu_skilled_head_dummy Occu_craft_head_dummy Occu_operator_head_dummy
  Occu_elementary_head_dummy urban_rural rural_urban Cornwall Middlesex Surrey popquint
  hhld_mems_in_sch hhld_mems_preprimary hhld_mems_primary hhld_mems_secondary hhld_mems_techvoc
  hhld_mems_tertiary hhld_mems_other share_health_insure chronic_count_per_hhld genhealth
  /MATCHIT
    MATCH=NEAREST
    EST =LOGIT
    DISCARD = BOTH
    MORDER = LARGEST
    RATIO = 1
    CALIPER = .0075
  /PLOT HISTPLOT JITTERPLOT HISTBAL DOTPLOT INDBAL RESOLUTION = 200
  /OUTPUT PAIRED MATCHED_CASES_ONLY.

***PROPENSITY SCORE MATCHING NEAREST NEIGHBOR WITHOUT REPLACEMENT WITH CALIPER 0F 0.005 using all 48
VARIABLES**
```

```
DATASET ACTIVATE DataSet1.
PSMATCHING3
  /VARS
     TREAT = HHOLDTYPE2
     COVS = hhsize2 BioDadPresent_dummy BioMomPresent_dummy Children boys girls child6_dum1
    child6_10_dum1 child11_17_dum1 males15 females15 working_age fem_working_age mal_working_age
    males65 females65 depend malhead_dum femhead_dum age_head marriedhead yrsofsch_head Occu_none_dummy
    Occu_mngr_head_dummy Occu_professional_head_dummy Occu_assc_prof_head_dummy Occu_clerks_head_dummy
    Occu_service_head_dummy Occu_skilled_head_dummy Occu_craft_head_dummy Occu_operator_head_dummy
    Occu_elementary_head_dummy urban_rural rural_urban Cornwall Middlesex Surrey popquint
    hhld_mems_in_sch hhld_mems_preprimary hhld_mems_primary hhld_mems_secondary hhld_mems_techvoc
    hhld_mems_tertiary hhld_mems_other share_health_insure chronic_count_per_hhld genhealth
  /MATCHIT
     MATCH=NEAREST
     EST =LOGIT
     DISCARD = BOTH
     MORDER = LARGEST
     RATIO = 1
     CALIPER = .005
  /PLOT HISTPLOT JITTERPLOT HISTBAL DOTPLOT INDBAL RESOLUTION = 200
  /OUTPUT PAIRED MATCHED_CASES_ONLY.
```

***PROPENSITY SCORE MATCHING NEAREST NEIGHBOR WITHOUT REPLACEMENT WITH CALIPER 0F 0.0025 using all 48 VARIABLES**

```
DATASET ACTIVATE DataSet1.
PSMATCHING3
  /VARS
     TREAT = HHOLDTYPE2
     COVS = hhsize2 BioDadPresent_dummy BioMomPresent_dummy Children boys girls child6_dum1
    child6_10_dum1 child11_17_dum1 males15 females15 working_age fem_working_age mal_working_age
    males65 females65 depend malhead_dum femhead_dum age_head marriedhead yrsofsch_head Occu_none_dummy
    Occu_mngr_head_dummy Occu_professional_head_dummy Occu_assc_prof_head_dummy Occu_clerks_head_dummy
    Occu_service_head_dummy Occu_skilled_head_dummy Occu_craft_head_dummy Occu_operator_head_dummy
    Occu_elementary_head_dummy urban_rural rural_urban Cornwall Middlesex Surrey popquint
    hhld_mems_in_sch hhld_mems_preprimary hhld_mems_primary hhld_mems_secondary hhld_mems_techvoc
    hhld_mems_tertiary hhld_mems_other share_health_insure chronic_count_per_hhld genhealth
  /MATCHIT
     MATCH=NEAREST
     EST =LOGIT
     DISCARD = BOTH
     MORDER = LARGEST
     RATIO = 1
     CALIPER = .0025
  /PLOT HISTPLOT JITTERPLOT HISTBAL DOTPLOT INDBAL RESOLUTION = 200
  /OUTPUT PAIRED MATCHED_CASES_ONLY.
```

***PROPENSITY SCORE MATCHING NEAREST NEIGHBOR WITHOUT REPLACEMENT WITH CALIPER 0F 0.001 using all 48 VARIABLES**

```
DATASET ACTIVATE DataSet1.
PSMATCHING3
  /VARS
     TREAT = HHOLDTYPE2
     COVS = hhsize2 BioDadPresent_dummy BioMomPresent_dummy Children boys girls child6_dum1
    child6_10_dum1 child11_17_dum1 males15 females15 working_age fem_working_age mal_working_age
    males65 females65 depend malhead_dum femhead_dum age_head marriedhead yrsofsch_head Occu_none_dummy
    Occu_mngr_head_dummy Occu_professional_head_dummy Occu_assc_prof_head_dummy Occu_clerks_head_dummy
    Occu_service_head_dummy Occu_skilled_head_dummy Occu_craft_head_dummy Occu_operator_head_dummy
    Occu_elementary_head_dummy urban_rural rural_urban Cornwall Middlesex Surrey popquint
    hhld_mems_in_sch hhld_mems_preprimary hhld_mems_primary hhld_mems_secondary hhld_mems_techvoc
    hhld_mems_tertiary hhld_mems_other share_health_insure chronic_count_per_hhld genhealth
  /MATCHIT
     MATCH=NEAREST
     EST =LOGIT
     DISCARD = BOTH
     MORDER = LARGEST
     RATIO = 1
     CALIPER = .001
  /PLOT HISTPLOT JITTERPLOT HISTBAL DOTPLOT INDBAL RESOLUTION = 200
  /OUTPUT PAIRED MATCHED_CASES_ONLY.
```

```
******
***SENSITIVITY ANALYSIS USING NEAREST NEIGHBOUR MATCHING WITH REPLACEMENT (1:5)  AND CALIPER OF 0.005
***
******

DATASET ACTIVATE DataSet1.
PSMATCHING3
  /VARS
     TREAT = HHOLDTYPE2
     COVS = hhsize2 BioDadPresent_dummy BioMomPresent_dummy Children boys girls child6_dum1
    child6_10_dum1 child11_17_dum1 males15 females15 working_age fem_working_age mal_working_age
    males65 females65 depend malhead_dum femhead_dum age_head marriedhead yrsofsch_head Occu_none_dummy
    Occu_mngr_head_dummy Occu_professional_head_dummy Occu_assc_prof_head_dummy Occu_clerks_head_dummy
    Occu_service_head_dummy Occu_skilled_head_dummy Occu_craft_head_dummy Occu_operator_head_dummy
    Occu_elementary_head_dummy urban_rural rural_urban Cornwall Middlesex Surrey popquint
    hhld_mems_in_sch hhld_mems_preprimary hhld_mems_primary hhld_mems_secondary hhld_mems_techvoc
    hhld_mems_tertiary hhld_mems_other share_health_insure chronic_count_per_hhld genhealth
  /MATCHIT
     MATCH=NEAREST
     EST =LOGIT
     DISCARD = NONE
     MORDER = LARGEST
     RATIO = 5
     CALIPER = .005
     REPLACE
  /PLOT HISTPLOT JITTERPLOT HISTBAL DOTPLOT INDBAL RESOLUTION = 200
  /OUTPUT PAIRED MATCHED_CASES_ONLY.
```