



Girls' Economic Empowerment – The Best Contraceptive?

*Results from a long-term pilot survey with focus on
entrepreneurial behavior*

Monica B-G Blaalid and Reidun Kristina Malvik

Supervisor: Kjetil Bjorvatn

Master Thesis in Economics

NORWEGIAN SCHOOL OF ECONOMICS

This thesis was written as a part of the Master of Science in Economics and Business Administration at NHH. Please note that neither the institution nor the examiners are responsible – through the approval of this thesis – for the theories and methods used, or results and conclusions drawn in this work.

Abstract

This paper is written as the final thesis of our master degree in Economics at the Norwegian School of Economics, NHH. The thesis is a contribution to the Girl Empowerment Project – a research project launched by the Christian Michelsen Institute (CMI), the Norwegian School of Economics (NHH), the Economic and Social Research Foundation (ESRF), Femina HIP and the Development Pioneer Consultants (DPC).

Teenage pregnancies are common in many low-income countries, but the reasons for why teenage girls become pregnant are not well understood. Given the health and economic downfalls that can come with early pregnancy, the Girl Empowerment Project investigates whether adolescent pregnancies can be countered by empowering young women. This is done by providing young women in Tanzania with two different empowerment strategies: an information treatment, containing information on reproductive health, gender equality and rights, and an opportunity treatment, providing the girls with entrepreneurship training to improve their skills and knowledge of how to run a business. By comparing the two treatments, and a combination of the two, the study provides insights about their relative importance of providing teenagers with information and opportunities.

As a part of the project, we have collected new data from a subsample of the girls, with updated information of where they are today. The process of collecting the data also contributes to the project by providing useful insights to the planning of the next and final round of data collection.

This thesis compares the new data to data that has been collected in previous surveys, and evaluates the effects of the treatments on girls' plans to start a business and income-generating activities. We find that the business training has successfully encouraged the girls to develop business plans, and findings from the new data suggest that business training also has inspired them to commence income-generating activities. The findings further indicate that the business training has made the girls better equipped to overcome constraints to starting and running a business.

Acknowledgements

First of all we would like to thank our supervisor Kjetil Bjorvatn for inviting us to take part in the Girl Empowerment Project. His guidance and enthusiasm have made this a rich and enjoyable learning experience for us.

This thesis would not have been the same without contributions from the entire research team of the Girl Empowerment Project. We would like to thank Bertil Tungodden and Lars Ivar Oppedal Berge, as well as Ranveig Falch and Amina Mohamed Maalim, for their recommendations and cooperation. We offer a special thanks to Vincent Somville for consistently taking the time to assist us with practical issues and offering suggestions to how we might improve our thesis.

We would also like to express gratitude to the team in Dar es Salaam for assisting us with the implementation of the pilot survey: Linda H. Sekei, Juda Lyamai, Ramlath Yassi, Swalehe Manture, Jacqueline Mgumia, and Femina HIP, as well as Tausi Kida and ESRF for hosting us. Thank you all for making us feel at home in Dar es Salaam and for teaching us about the local culture.

Norwegian School of Economics

Bergen, 20.06.2014

Monica Backer-Grøndahl Blaalid

Reidun Kristina Malvik

Table of Contents

1 Introduction	6
1.1 Entrepreneurship – the best contraceptive?	6
1.2 The Girl Empowerment Project	7
1.3 Research Question	9
1.4 Outline	10
2 Literature review	11
3 Institutional framework	15
3.1 Key facts about Tanzania	15
3.2 Education in Tanzania	16
3.3 Small-scale entrepreneurship in Tanzania	18
4 Experimental design	20
4.1. Randomized Controlled Trials	20
4.2 Sample selection	23
4.3 Timeline of the project	25
5 Analysis	29
5.1 Data	29
5.2 Econometric specifications	33
6 Call history from the long-term pilot survey	35
7 Findings	40
7.1 Findings from phone interviews in LTPS	40
7.2 Treatment effects on having plans to start a business	49
7.3 Treatment effects on having income from work	56
8 Conclusion	64
References	65
List of Figures	71
List of Tables	71
Abbreviations	72
Appendix A: Interview Guide for Phone Interviews LTP	73

Appendix B: Probit estimations	78
Table B1: Plans to start a business (STS).....	78
Table B2: Income (STS)	79
Appendix C: Marriage, moving and treatment effects	80
Table C1: Marriage and permanent relationships (LTPS)	80
Table C2: Moving (LTPS)	81
Appendix D: Business knowledge questions from baseline survey.....	82

1 Introduction

1.1 Entrepreneurship – the best contraceptive?

“When a girl becomes pregnant, her present and future change radically, and rarely for the better. Her education may end, her job prospects evaporate, and her vulnerability to poverty, exclusion and dependency multiply.”

UNFPA: *State of World Population 2013: Motherhood in Childhood*, (2013, page 2)

Every year more than 7 million girls under the age of 18 give birth (UNFPA, 2013). While the number of pregnancies in developing countries is declining, as many as 95 percent of the world’s births to adolescents occur in the developing world (UNFPA, 2013). High poverty rates, inadequate health care and reduced social welfare make teenage pregnancies comparatively more problematic in developing countries than in industrialized countries. Complications of pregnancy and childbirth are the leading cause of death among women aged 15-19 years in low and middle-income countries (WHO, 2012), and stillbirths and newborn deaths are 50 percent higher among babies born to mothers who are younger than 20 years of age, when compared with mothers who are 20-29 years of age (WHO, 2012).

In addition to the health risks associated with adolescent pregnancy, there can be severe societal consequences for a young mother. A pregnancy may lead to school dropout, which can, in turn, jeopardize the girl’s future economic prospects and exclude her from other opportunities (UNFPA, 2013). As possible job prospects evaporate, she may be forced into poverty, and may become dependent on others to look after her and her child. It follows from these findings that postponing pregnancy can have positive effects on health outcomes, educational attainment (Goldin & Katz, 2000; Rasul, 2008) and future employment of women (Bailey, 2006; Miller, 2010).

Although the consequences of teenage pregnancies may seem obvious, teenage pregnancies have proved difficult to prevent, as identifying the underlying causes has proven a challenging task. There may be circumstances beyond the girl’s control and the girl’s absence of choices that partly cause adolescent pregnancy. In other words, it might be the girl’s choice to start a family in the sense that she is not being forced to do so, but her lack of alternatives may mean that it was not a choice made of free will.

Societal context plays a central role in contributing to teenage pregnancy through attitudes towards marriage and fertility, gender norms, and female violence. Lack of knowledge about sex, contraception and family planning is another underlying cause of teenage pregnancies. Many informational campaigns have been implemented to prevent adolescent pregnancies and transmission of diseases such as HIV/AIDS (WHO, 2012). However, global measures on sexuality education show that only a fourth of young women in low- and middle-income countries have comprehensive and correct knowledge of how to prevent the spread of HIV (WHO). This indicates that there is still much work to be done regarding the sexual education of young women and around the world.

There is a clear link between teenage pregnancies and education – higher birth rates are observed among women with lower levels of education than among women who have received secondary or tertiary education (National Bureau of Statistics, 2011). Regional context and marital status are other influencing factors. Girls from poor backgrounds and limited access to education tend to marry at a younger age, and early marriage leads to early pregnancy (Muthengi-Karei & Erulkar, 2012). In Tanzania, this trend can be observed regionally. In rural areas, it is more common for young women to marry early, with the median age at first marriage being 18.5 years, compared to 19.8 years in urban areas in Tanzania (National Bureau of Statistics, 2011). As a result of many premature marriages, young childbearing is more common in rural areas.

In countries like Tanzania, young women, with limited prospects of further education or employment, may consider marrying and having children as their best option. However, entrepreneurship training can act as a form of contraception by giving the girls the necessities to provide for themselves and their immediate family.

1.2 The Girl Empowerment Project

The United Nations Populations Fund (UNFPA) recommends multidimensional strategies towards empowering girls in order to prevent adolescent pregnancies: *“Many actions by governments and civil society that have reduced adolescent fertility were designed to achieve other objectives, such as keeping girls in school, preventing HIV infection, or ending child marriage”* (UNFPA, 2013).

This is what the Girl Empowerment Project has set out to do. The Christian Michelsen Institute (CMI) has, in cooperation with Norwegian School of Economics (NHH), Femina HIP, the Economic and Social Research Foundation (ESRF) and the Development Pioneer Consultants (DPC), developed a project called “Girls’ Economic Empowerment – The Best Contraceptive?” as a contribution to the existing literature in the field. Their main research objective has been to seek a better understanding of why the majority of young girls decide to start childbearing at an early age, and how this interacts with and affects their economic situation (Tungodden, 2012).

The project was conducted with girls in their last four months of secondary school, when they are on the verge of making two of the most important decisions in their lives: what to do when leaving school and whether to start childbearing. By using the randomized controlled trial methodology, two different treatments are offered to a randomly selected sample of girls: an information treatment and an opportunity treatment. The information treatment provides practical information about gender equality and reproductive health. The opportunity treatment provides the girls with knowledge on how to run their own business with the aim of expanding the girls’ economic opportunities. By comparing outcomes from the two treatment interventions and an intervention that combines the two, the study hopes to identify the most effective intervention in changing the fertility and economic aspirations of young women, and to what extent a combination of the two could have an “additional” effect beyond giving the girls both treatments (Tungodden, 2012). The three interventions were introduced to a total of 3,485 girls distributed across 80 schools in Tanzania.

Health information campaigns have been widely used to prevent risky behavior, such as adolescent pregnancy. Existing research has shown that households in developing countries are responsive to information, but that they often lack information on the returns of investing in preventive behavior or information on prevention techniques (Dupas, 2011a). A randomized controlled trial on sexual behavior in Kenya found that teenagers felt more empowered and reduced their risky sexual behavior when provided with risk information (Dupas, 2011b). In general, the underlying idea of information treatment is that adolescent pregnancy is caused by lack of relevant information and personal control. Despite endless information campaigns about the risks related to adolescent pregnancy, it continues to be a widespread problem. The Girl Empowerment Project questions whether adolescent

pregnancies reflect lack of economic opportunities, and whether young women need an alternative to motherhood in order to change their fertility and economic aspirations.

By providing the girls with the necessary knowledge and skill to start a business, entrepreneurship is promoted as an alternative future path to starting a family. Lack of human capital is increasingly being recognized as a constraint to the development of small businesses. Although the evidence is somewhat mixed, there is comprehensive literature supporting the importance of business training in fostering entrepreneurship and improving business outcomes (de Mel, McKenzie, & Woodruff, 2012b; Klinger & Schündeln, 2007). Findings suggest that the new businesses started after the training can be expected to have higher profitability and better practices, than those started by untrained entrepreneurs (de Mel, McKenzie, & Woodruff, 2012b). To summarize, the literature indicates that providing entrepreneurship training to young women in developing countries may empower them and improve their economic prospects, which further could make them postpone plans to start a family.

This paper is a contribution to the Girl Empowerment Project. Based on the data collected at a baseline survey in April-May 2013 and a short-term survey in September-October 2013, we analyze the effect of the various treatments on the girls' business plans. In April-May 2014, we conducted a long-term pilot survey where we interviewed 179 girls from four of the schools, each representing a treatment group. There were two main reasons for conducting this long-term pilot survey: to collect updated data on what the girls are doing today, and to map the process of reaching the girls. The updated information about the girls' current plans to start a business and their income-generating activities is the main focus of this paper. Information regarding the phone interview process provides useful information for the planning of the long-term survey that is expected in September-October 2014.

As this paper focuses on plans to start a business and the entrepreneurship aspects of the Girl Empowerment Project, there will be limited, if any, focus on the health aspects of the project.

1.3 Research Question

Based on a sample composed of 179 girls from four schools, with each school representing a treatment group, we attempt to answer the following research question:

What are the daily activities of girls who have recently graduated from secondary school – do they have plans to start a business and do they have a regular income? Do we observe any systematic differences across treatment groups in regards to their business plans, income-generating activities and the constraints that the girls are facing?

The results of this analysis shed light on entrepreneurship among young women in developing countries, and demonstrate the role business training can play in inspiring and providing young women with the skills needed to start a business.

1.4 Outline

The rest of the study is organized as follows: Chapter 2 presents an overview of relevant literature. In Chapter 3, we provide some context to the analysis. Chapter 4 presents the design of the project, while Chapter 5 explains how data was collected and analyzed. Findings from the interview process of the long-term pilot survey are presented in Chapter 6, while other results are presented and analyzed in Chapter 7. We conclude in Chapter 8.

2 Literature review

Current and potential small-scale entrepreneurs in developing countries face a variety of obstacles in their daily business operations. One significant barrier is capital constraints, which makes it difficult for entrepreneurs to build assets and manage working capital. Traditionally formal banks have not considered poor people in their customer portfolios, leaving the poor to use informal financial instruments, which tend to be costly, and unreliable (Collins, Murdoch, Rutherford, & Ruthven, 2009). The poor's limited access to financial services resulted in the rise of the microfinance movement. Those promoting microfinance intended to empower small-scale entrepreneurs, particularly in rural areas, by providing them access to finance in the hopes that this would lift them out of poverty (IMF, n.d.).

However, economic field experiments show that the impact of financial capital, either in the form of credit or grants, depends on other dimensions in order to lead to business growth (Banerjee, Duflo, Glennester, & Kinnan, 2013; de Mel, McKenzie, & Woodruff, 2008; Karlan & Morduch, 2009). Individual characteristics, such as limitations to human capital, is one of several obstacles faced by poor entrepreneurs, and are likely to be crucial determinants of entrepreneurial success. Individual characteristics are not typically addressed by microfinance institutions, but following the mixed evidence on the importance of financial constraints, attention has shifted towards lack of human capital and other constraints to microenterprise development.

Business training may improve entrepreneurs' level of business knowledge and skills, enabling them to make better strategic decisions. By increasing their human capital, business training may help entrepreneurs to deal more effectively with obstacles such as capital constraints, demand deficiency, and poor infrastructure. There is now comprehensive literature on the role of business training in fostering entrepreneurship and improving business outcomes. Klinger and Schündeln (2007) find that participation in business training significantly increases the probability of establishing new enterprises and expanding existing businesses. However, research points to a stronger effect of training for male entrepreneurs, while there are limited or non-existent treatment effect for females (Berge, Bjorvatn, & Tungodden, 2014; Gine & Mansuri, 2011; Klinger & Schündeln, 2007).

Berge, Bjorvatn and Tungodden (2014) conducted a field experiment in Tanzania offering a combination of business training and financial grants to entrepreneurs who were already

members of the largest microfinance institution in Tanzania, Promotion of Rural Initiatives and Development Enterprise (PRIDE). Their study concluded that business training in combination with a business grant caused a substantial improvement in profits for male entrepreneurs, but had very limited impact on business outcomes for female entrepreneurs. Interestingly, the authors found that business training improved the business knowledge of both male and female participants, as well as changing their mindset. The authors suggest that both gender differences in mind-sets, particularly regarding willingness to compete, and gender differences in social constraints, may explain the gender difference in treatment effects on business outcomes.

Gine and Mansuri (2011) conducted a field experiment for both women and men in rural Pakistan, offering microfinance clients training and the possibility to access a business loan. They found that the business training improved business practices, increased business knowledge and improved several household outcomes for men, while women improved their business knowledge after the training, but show no improvements in the other outcomes. The authors point out that 40 percent of the businesswomen reported that their spouses were responsible for most of their business decisions. Thus, the limited decision making control women have in their own businesses can be seen as a reason for why female businesses show no improvement after business training.

While most of the articles written in the field explore treatment effects on established small-scale entrepreneurs, this thesis looks at young women with no established enterprise prior to the training. Research focusing on young women in this context is limited, though a few papers take age into consideration. Bruhn and Zia (2011) focus on young borrowers of business loans in Bosnia and Herzegovina with clients aged 18 to 35. Bosnia is an emerging economy that is struggling with high youth unemployment and low business survival. The authors studied the effects of business and finance training on existing business owners, as well as on potential entrepreneurs. The age of the client was not found to predict whether the client would be interested in participating in the business and financial education course. Bruhn and Zia found no significant treatment effect on business start-ups, suggesting that business training alone is not enough to promote growth of new businesses in emerging markets.

De Mel, McKenzie and Woodruff (2012b) studied the impact of business training, as well as the combined effect of training and a cash grant, on women in urban Sri Lanka. They looked

at two groups of women: women who were already operating enterprises and women who were interested in starting a business but were currently out of the labor force. They found that among women who did not have a business prior to the intervention, the profitability of newly established businesses improved from the business training. The women who already had a business at the time of the intervention saw no impact from the business training alone on their business outcome. Hence, the training sped up the entry for women interested in starting a business, but did not lead to an increase in net business ownership. For new entrants, the training increases both profitability and business practices. This suggests that the business training may be more effective for new business owners than for those whose enterprises exist prior to the training.

Another paper that is highly relevant for the Girl Empowerment Project is the Ruka Juu study by Bjorvatn, Cappelen, Sekei, Sørensen and Tungodden (2014). The paper is based on the reality show and entrepreneurship competition for youth in Tanzania called Ruka Juu (meaning: “Jump Up”) which aired on national television. In a country where high joblessness rates force youth to create their own jobs, Ruka Juu is meant to educate and inspire youth to act and start their own businesses. The study is novel in that it evaluates the impact of education and inspiration through television, focusing on youth, and in particular young women. The study found that Ruka Juu inspired the viewers and led them to adopt a more entrepreneurial mind-set, with treatment effects being particularly pronounced among female viewers.

The literature review shows that human capital interventions are often less effective for female entrepreneurs than for their male counterparts. In various research articles different arguments are raised for why this is the case. The girls’ low willingness to compete, in combination with the characteristics of their mind-sets, is suggested as a possible explanation for the gender effect, as women have limited control of their business decisions in some contexts. However, these studies are both based on women who had businesses prior to receiving business training. Other papers suggest that business training is more effective for new business owners than for those whose enterprise existed prior to the training. Business training can play an important part in informing and inspiring youth to start a business, as well as influence their way of thinking, in order to make them believe they can succeed as entrepreneurs.

The research also indicates that business training by itself may not be enough to promote growth of new businesses in emerging markets. Small-scale entrepreneurs face multiple and complex constraints and one cannot expect all of these to evaporate as a result of the business training. However, the findings still indicate that business training may enable entrepreneurs to make better decisions for their business and better deal with other barriers to success.

Studies focusing on youth who does not have a business prior to the intervention, is limited. This paper aims at providing insights to the effectiveness of business training on young women's plans to start a business, and the implementation of their plans.

3 Institutional framework

The following chapter highlights characteristics of the Tanzanian society, such as economic development, the educational situation, and adolescent pregnancies. These aspects are all relevant when analyzing findings on young Tanzanian women and their plans for the future. Of particular interest are the features of small-scale entrepreneurship in the country.

3.1 Key facts about Tanzania

The United Republic of Tanzania was constituted in 1964, when the two former British colonies Tanganyika and Zanzibar, merged after achieving their independence in the early 1960s. The country is located in Eastern Africa and has a population of about 48 million. The political capital of Tanzania is Dodoma, while the main commercial city is Dar es Salaam. Like many developing countries Tanzania has a young population – the median age is 17.3 years, and 45 percent of the population are children under the age of 15 (CIA, 2013). Tanzania has one of the 20 fastest growing populations in the world, with a fertility rate of five children born per woman and a life expectancy of 60.8 years (CIA, 2013). The young population represents an opportunity for the country's prosperity, but presents a challenge for Tanzania, considering the low job creation in the formal sector compared to the growth in the labor force.

Over the past ten years Tanzania's economy has been growing steadily with a GDP growth rate of six to seven percent per year in 2009-2013. Despite the growth of the economy, the country remains one of the world's poorest economies in terms of income per capita. It ranks as 152 on the Human Development Index, and the country's Gini coefficient¹ is 37.6. Tanzania is still predominantly rural with about 75 percent of the population living in rural areas where poverty is common (World Bank, 2013). The country's progress within poverty reduction and human development has so far been slow compared to the growth the country

¹ The Gini coefficient is the most common method for comparing income differences between countries. A Gini coefficient of 0 indicates that the income is evenly spread among the citizens of the country, while a Gini coefficient of 100 indicates that one person receives all the income. The Tanzanian Gini coefficient of 37.6 (2007) indicates that there are some income variations in the population. As a comparison, Norway has a Gini coefficient of 25 (2008), while the South African coefficient is 63.1 (2009). Hence, Tanzania has relatively small income differences compared to countries like South Africa, but still more variations than what is observed in Norway.

has experienced. Tanzania is striving to become a middle income country within 2025 through the Tanzania Development Vision 2025. This emphasizes the country's goal of becoming a prosperous nation through eradicating poverty, ignorance and disease (TradeMark Southern Africa, 2013). Several areas are identified that Tanzania needs to address in order to achieve status as a middle income country, involving necessary improvements within education and the quality of learning.

3.2 Education in Tanzania

The educational system of Tanzania is based on primary, secondary and tertiary levels. The seven-year primary schooling is compulsory and tuition free, but parents still have to pay testing fees, academic contributions, and more (Helgesson, 2006). The students must pass the Primary School Leaving Examination (PSLE) at the end of their seventh year in primary school to be eligible to attend public secondary school. The secondary level is divided into the ordinary levels of Forms I to IV, and the advanced levels of Form V to VI. At the end of Form IV, all qualified students are allowed to sit for the Certificate of Secondary Education Examinations (CSEE). The examinations take place in November and tests students in nine different subjects. After the examination students are awarded with division I, II, III, IV or 0, depending on to what extent they meet the requirements (NECTA). Only students passing at least five of nine examinations with grades A to C may progress to the advanced levels (NECTA). Students that are awarded division 0, fail. In order to receive the exam results, one has to pay an examination fee of 50,000 Tanzanian shillings (TZS) (NECTA). With late payment (after February 31st) comes an additional fee of 15,000 TZS, while the student must pay a further 35,000 TZS if fees are not paid before the results are published in April.² Thousands of students do not receive their exam results every year because they are unable to pay the fees. Among all students who wrote their Form IV examinations (CSEE) in 2013, eight percent³ have yet to receive their results due to unpaid fees (Daily News, 2014).

² The examination fees for the CSEE exam were recently increased, and the data on the webpage of NECTA is not up to date. We contacted the principal at one of the schools that participated in the long-term pilot survey to confirm the current examination fees. Fees listed in this text are based on the information this principal gave us.

³ From the 2013 examinations of Form IV, 31 518 out of the 404 083 students sitting for the exam has not received their results because they have not paid their examination fees.

The passing rates for the final examination in Form IV are low all over Tanzania. In 2013, 58 percent of the students who sat for the National Form IV examinations passed. Divisions I, II or III were rewarded to 21 percent of the students who sat for the exam. Even though seven girls made it to the top ten in the ranking of candidates from the 2013 November exam (IPP Media, 2014), girls constituted only 37 percent of the students achieving the three highest divisions (IPP Media, 2014). This indicates that there is a strong gender difference in the academic achievements of Tanzanian secondary students.

Tanzania has made significant progress in improving access to primary education, which makes it seem like Tanzania is well on track to achieve the Millennium Development Goal (MDG) of universal access to primary education by 2015 (UN, 2013). The abolishment of school fees in primary schools in 2001, along with a compulsory requirement that parents send all children to school, are two reasons why 94 percent of school aged children were enrolled in primary school in 2011 compared with 59 percent in 2000. Secondary school enrollment has also expanded from six percent in 2000 to 30 percent in 2011 (Ministry of Education and Vocational Training, 2012).

However, poor quality of education remains a major challenge for the government, partly because the enrollment expansion was not met with the necessary supply of qualified teachers.⁴ Approximately one-third of the children are failing to complete the primary level and less than 60 percent passed the PSLE at the end of 2011. As the target of the MDG of universal access to primary education is to ensure that all children will be able to *complete a full course of primary schooling* (UN, 2013), the poor quality of education may prevent Tanzania from reaching the goal. The low quality of education in the primary level has also affected the quality of secondary education (World Bank, 2013).

Although girls' enrollment is close to parity with boys' enrollment at all primary education levels, there are still gender disparities in completion rates and performance. After the age of 14, girls are less likely than boys to be in school. Cultural beliefs and customs, early pregnancy and marriage are among the key factors for the gender disparities in retention and completion (US Aid, 2013). Girls persistently perform worse in national school examinations

⁴ The Tanzanian Five Year Development Plan for 2011/2012-2015/2016 identify that the country needs to have more than 900,000 qualified teachers employed in order for Tanzania to realize the status as a middle income country. The current level is 238,000 teachers, which illustrates the current situation of the country's educational system (Government of Tanzania, 2011).

at both primary and secondary levels, especially in mathematics and science (UN Tanzania, n.d.).

With regards to primary education there are also regional disparities, and uneven distribution of education expenditure persist (World Bank, 2013). Poverty is a widespread phenomenon in rural Tanzania and is reflected in lower quality and poorer access to education. As a result, educational outcomes are consistently higher in urban than in rural areas (UNICEF, n.d.). It is evident that in Tanzania, educational level and performance varies by gender, region and socioeconomic status (US Aid, 2013).

3.3 Small-scale entrepreneurship in Tanzania

As in most countries in Sub Saharan Africa, neither the public sector nor the private sector is able to provide enough jobs for the ever-expanding labor force. As school fees and academic requirements place higher education out of reach for most Tanzanian youth, many of them are faced with limited opportunities to get a job after finishing their basic education.

In developing countries, microenterprises employ more than half of the labor force, and developing and establishing more enterprises is a key policy concern in many of these countries (Berge, Bjorvatn, & Tungodden, 2014). Leonard and Mfaume (2004) argue that most small-scale entrepreneurs have chosen self-employment in the *informal* sector as a last resort to earn an income. A majority of Tanzania's micro and small enterprises are informal, which implies that they are not registered with the government and do not pay tax. The informal sector consists mainly of small-scale traders, farmers, small manufacturers, craftsmen, individual professionals and small-scale businesses (ESRF, 2012). The number of informal enterprises in Tanzania is growing fast and the sector is contributing about 40 per cent of the country's GDP. Many microenterprises have no paid employees (TCCIA, 2009). The informal sector is increasingly becoming recognized as a mechanism of economic adjustment and source of livelihood for the poor and unemployed (Hobson, 2011).

Djankov, Lieberman, Mukherjee, and Nenova (2002) make the distinction between "subsistence enterprises" and "unofficial enterprises", based on their potential to move into the formal sector.⁵ They argue that young entrepreneurs are likely to start out with a

⁵ The formal sector encompasses all jobs with normal hours and regular wages, and is recognized as income sources on which income taxes must be paid. (Business Dictionary, n.d.)

subsistence enterprise. These enterprises are characterized as possessing low potential and experiencing higher failure rates, as they are often labor intensive and sell highly homogenous products. Unofficial enterprises have relatively high potential, are more dynamic and are likely to be formally registered, although they may engage in some unofficial activities or using unreported employees. Subsistence enterprises include street traders, micro-enterprises and subsistence farmers, and a substantial segment of the entrepreneurs are poor, with low education and skill levels (Ishengoma & Kappel, 2006).

4 Experimental design

This chapter presents the experimental design used in the Girl Empowerment Project, and more specifically, in the long-term pilot survey (LTPS). The first section discusses Randomized Controlled Trials (RCTs) as the chosen research design for the project. In section two we describe the sample selection process. In the third section we present the timeline of the project.

4.1. Randomized Controlled Trials

In policy work, interventions are often implemented to solve problems or inefficiencies in a society. Interventions may include anything from supplying books to schools in order to increase learning outcomes, to giving out free mosquito nets to reduce the spread of malaria. However, one must be careful when it comes to evaluating the true benefits of interventions. If, for example, one happened to observe a reduction in the spread of malaria after the distribution of mosquito nets, other factors might have caused the reduction. In order to claim that an intervention has had an effect, it is necessary to compare the same group with and without the treatment intervention. Since the two alternatives are mutually exclusive, it is impossible to observe the exact same sample in both situations.

When RCTs are designed and implemented correctly, they provide an unbiased and internally valid estimate of the average treatment effect. For this reason, many considers RCTs to be the ‘gold standard’ for determining causal relationships (Drageset & Ellingsen, 2009). Randomization is central to the method, and requires randomly assigning the participants to two groups; a treatment group and a control group. The treatment group receives an intervention, while the control group receives nothing or a placebo. As a result of the randomization, the two groups are in theory statistically equivalent to each other and to the sample population as a whole (Poverty Action Lab, n.d.). This means that they are identical on both observable and unobservable characteristics. However, this requires that the sample is large enough to detect meaningful effects of the intervention. With this in place, one can state that an effect from the treatment is caused by the intervention.

This reasoning can be shown formally through the Holland-Rubin framework presented by Deaton (2010). The framework states that each individual i in the population is randomly

assigned either to the treatment $T_i = 1$, or to the control group $T_i = 0$. The outcomes associated with each individual are Y_{i0} or Y_{i1} , depending on whether the individuals receive the treatment or not. The aim is to identify the size of the treatment effect $Y_{i1} - Y_{i0}$. The problem is that the individuals are assigned to either the treatment or the control group. As a result, one cannot observe both outcomes. However, it is possible to obtain the difference between the average outcomes of the treatment group and the average outcomes of the control group. But such a comparison does not provide the answer to what we are looking for, which is the causal effect of the treatment. Formally, the comparison of average outcomes for the treated and the non-treated is linked to the average causal effect by the following equation:

(4.1.1a)

$$\begin{aligned} E(Y_{i1}|T_i = 1) - E(Y_{i0}|T_i = 0) \\ = [E(Y_{i1}|T = 1_i) - E(Y_{i0}|T = 1_i)] + [E(Y_{i0}|T_i = 1) - E(Y_{i0}|T_i = 0)] \end{aligned}$$

The bracket on the left side is the observed difference on average outcome, while the first bracket on the right side is the average treatment effect on the treated, which is the effect we are trying to identify. It captures the average difference between the outcome of the treated, $E(Y_{i1}|T_i = 1)$, and what would have happened to them had they not been treated, $E(Y_{i0}|T_i = 1)$. However, the observed difference in outcomes cannot separate the treatment effect from the selection bias, which is the second square bracket on the right side. It reflects systematic differences between the treatment group and the control group. If there are any non-treatment differences between the two groups, for example if the treatment group was characterized by more entrepreneurial experience or better access to capital prior to the treatment, making the selection bias positive. Since $E(Y_{i0}|T_i = 1)$ cannot be observed it is not possible to calculate the magnitude of the selection bias. As a result, the selection bias cannot be separated from the treatment effect, and one cannot state to what extent the treatment caused the difference in outcome between the treatment group and the control group (Angrist & Pischke, 2008).

Random assignment of individuals to control and treatment group eliminates the selection problem (Angrist & Pischke, 2008). As the expectation for randomized groups are identical on non-treatment characteristics, the selection bias will equal to zero. In other words, the selection bias converges towards zero as the sample size increases (Duflo, Glennerster, & Kremer, 2006). As the selection bias is removed by randomization, we are left with;

$$(4.1.1b) \quad E(Y_{i1}|T_i = 1) - E(Y_{i0}|T_i = 0) = [E(Y_{i1}|T_i = 1) - E(Y_{i0}|T_i = 1)]$$

This shows that the difference between the average treated outcome and the average untreated outcome equals the difference in the two observable outcomes. This gives us the difference of averages, while it is rather the average of the difference that is of interest. As the expectation is a linear operator, the difference of the average is identical to the average of the differences (Deaton, 2010). It can be rewritten as:

$$(4.1.1c) \quad E(Y_{i1}|T_i = 1) - E(Y_{i0}|T_i = 0) = E(Y_{i1} - Y_{i0}|T_i = 1)$$

This shows that the difference in means between the treatment group and control group is an estimate of the average treatment effect. This result is dependent on randomization, as well as the linearity of expectations.

Randomization can be done at individual or group level, depending on what is most beneficial. Spillover effects play an essential role when choosing the level of randomization. If the treatment group has spillovers to the control group, it can bias the estimation of the treatment effect. When data that has been randomized at the group level is analyzed at an individual level, it is important to note that the error term may not be independent across individuals, but correlated within the groups (Duflo, Glennerster, & Kremer, 2006). This can be explained as individuals in the same group are subject to the same shocks, making their outcomes correlated. The treatment effect is also uniform within groups, and as a result the correlation in outcomes can mistakenly be interpreted as an effect of the program (Duflo, Glennerster, & Kremer, 2006). As it is impossible to distinguish the correlated shock from the treatment effect, it is important to control for this when calculating the estimates, if one is to get correct standard errors. By clustering at the school level, we control for the common shocks the groups might have been subject to, and get more accurate standard errors for the estimation of treatment effects.

The methodology behind RCTs has met some criticism. Angus Deaton (2010) criticizes the use of randomized field studies and claims that the methodological problems with these studies are not recognized by researchers using RCTs. The technique obtains estimates for the average treatment effect of a program, but does not identify whether the distribution effect was wide or limited to a few special cases. Deaton (2010) argues that this leads to a focus that is too narrow and too local to identify what works in development, thereby making it difficult to design policies or to advance scientific knowledge about development processes based on

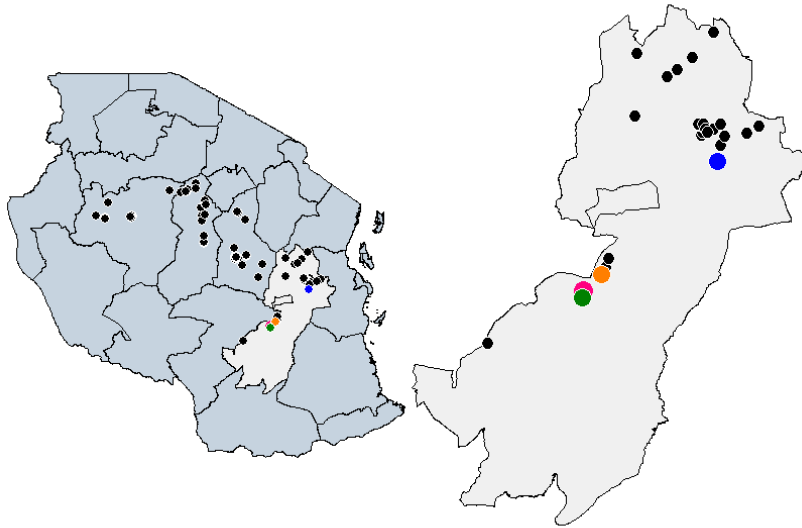
RCTs. The success of a project depends heavily on the cultural and social context; hence a project that is successful in one location may fail in another location. Deaton (2010) suggests that the analysis of projects refocus towards the investigation of potentially generalizable mechanisms that explain why and in what context the projects can be expected to work.

4.2 Sample selection

The Girl Empowerment Project uses the RCT methodology to investigate the causal effects of two empowerment strategies: an information treatment and an opportunity treatment. The information treatment contained information on reproductive health, gender equality, and rights. The opportunity treatment provided the girls with entrepreneurship training in order to improve their skills and knowledge of how to run a business. Further in the analysis we will refer to the information treatment as the *health* treatment, and the opportunity treatment as the *business* treatment. A third treatment group received a combination of the two treatments to investigate whether providing the girls with both treatments could have an “additional” effect beyond the fact that the girls received both treatments. A fourth group was assigned to the control group and left untreated. By comparing the four groups, the study provides insights about the relative importance of providing teenagers with information and opportunities.

The sample selection for the Girl Empowerment Project was done through random allocation of treatment to the schools in the survey. The sample included girls in their last year of secondary school in semi-urban areas, as this target group is thought to benefit the most from the interventions (Berge, Bjorvatn, Somville, & Tungodden, 2012). Most of the girls in this grade are in the age group 16 to 18 – an age when fertility increases drastically. When girls graduate from Form IV, they must decide what to do next. Given the low chances of being accepted for further studies or employment, their choice is often whether to start family or to start a business.

To be able to state that any causal relationship that we observe is generalizable for the whole population, the sample has to be representative of the total population. Schools with at least 20 girls in Form IV were sampled from the regions Dodoma, Morogoro, Singida and Tabora. 80 schools were selected to take part in the baseline survey to ensure a sufficient sample size. After the baseline survey was conducted, schools were randomly allocated to one of the three treatment groups or to the control group.

Figure 1: Distribution of schools

Note: The four colored dots represent the four schools in the long-term pilot survey (LTPS): the school in pink received health treatment, the blue school received business treatment, the green school received the combined treatment, and the orange did not receive any training. The map was made using the schools' GPS coordinates and graphing them onto a map in Stata.

The sample for the LTPS consisted of four schools in the rural areas of the Morogoro district. Their geographical location is shown in Figure 1. The four schools represent all treatment groups, and are the same four schools that Amina Mohamed Maalim⁶ focuses on in her master thesis.

Some of the girls still interact on a daily basis even though they are out of school, so by telephoning all girls from the four schools, we allowed “word of mouth” to play a role in the

⁶ Maalim is a master student at NHH and has completed her master thesis which can be viewed as an accompaniment to this one. Both of the theses are based on data from the baseline survey and short-term follow-up of the Girl Empowerment Project. Although they both focus on the business aspect of the project, the approaches are somewhat different; Maalim's thesis is largely based on personal essays written by the girls as a part of the short-term survey. By translating them from Kiswahili to English and then further analyzing the content, Maalim provides an in depth analysis of the girls' business plans and empowerment. The focus of this thesis, on the other hand, is on business practices. The two theses complement each other by providing the reader with a detailed view of how the programs are affecting realities on the ground in the somewhat longer term.

process of reaching the girls. This also made the process more similar to the work that will be done in the long-term survey, where the aim is to reach all girls.

Although findings from the pilot can enrich our knowledge and understanding of the girls who are in this phase of their lives, the four school sample does not fulfill the guidelines for achieving a representative sample⁷ and thus cannot be generalized to the full sample. Prior to the pilot we discussed the possibility of randomly selecting 200 girls – 50 from each treatment group. Even with a small sample, it could be possible to use the findings from a randomly selected sample to predict the status for the full sample. However, there was a concern that “word of mouth” could make the girls feel excluded when not being called, and could potentially cause extra work for the surveyors if girls started contacting them. Fears of “destroying” the data for later research made us choose the four previously mentioned schools for the pilot sample.

4.3 Timeline of the project

Table 1 shows the timeline of the Girl Empowerment Project. A baseline survey was conducted in March and April 2013. From June to September the same year, a training intervention took place, consisting of eight classroom sessions. Approximately five weeks after the training ended, a survey was conducted to measure the short-term impacts of the training interventions. As part of our master thesis, we conducted a LTPS on a subsample of the girls in April and May 2014. The contributions from the LTPS are updated data on where the girls are today, their plans to start a business and their income-generating activities, as well as insights regarding the process of reaching the girls. The findings will also be useful for the preparations of the more comprehensive long-term survey that is expected in September and October 2014.

The baseline survey and the short-term survey (STS) were both conducted by the Girl Empowerment Team using questionnaires. The baseline survey consisted of four thematic sections. The first section of the questionnaire looked at contact information as well as background information about the girls and their households. The second section covered personal competences by asking the girls their perceptions of time and risk, as well as their views of themselves and their future. This section also asked whether they had plans to start a

⁷ We reached and interviewed 179 girls from four schools representing all treatment groups.

Table 1: Timeline of the project

Name of activity	Type of survey	Month	Year
1: Baseline survey	Paper-based questionnaires	March-April	2013
2: Interventions	Training at the schools	June-September	2013
3: Short-term survey (STS)	Paper-based questionnaires	September-October	2013
4: Long-term pilot survey (LTPS)	Phone interviews	April-May	2014
5: <i>Future work: Long-term survey</i>	<i>Phone interviews</i>	<i>Expected: September-October</i>	<i>2014</i>

Note: (1)-(3) + (5) Number of observations: 3,485; (4) Number of observations 209.

business. The third section asked questions related to business and money, and the final section looked at the girls' opinions regarding marriage and family, as well as questions related to sexual health. The STS was more comprehensive than the baseline study and consisted of 13 thematic sections. The purpose of this survey was to measure short-term effects from the interventions. The long-term survey will identify the long-term effects of the interventions.

There are several advantages of using questionnaires compared to other methods for data collection. Questionnaires allow large amounts of information to be collected from a large number of people in a short period of time and in a cost effective way (Popper, 1959). The standardized nature of questionnaires limits the possibility of unreliability and invalidity of the survey. However, questionnaires have certain limitations. The standardized questions in questionnaires may be read differently by the respondents, and hence the replies will depend on how the respondent interpreted the question. As a result, the level of subjectivity may not be acknowledged. It is impossible to know how honest the respondent is being when answering the questionnaire, as well as how much thought the respondents put into their answers (Ackroyd & Hughes, 1981).

When we conducted the LTPS in April-May 2014, we made use of another data collection method. The pilot was designed as a phone interview, targeting 209 of the girls from STS. The interview guide for the phone interviews consisted of four thematically different

sections.⁸ The first section asked the girls to verify some personal information to update possible changes from the previous surveys. The second section investigated the extent to which girls had any source of regular income. In the third section, the girls were asked about their plans to start a business, and in the fourth section we asked them about constraints to entrepreneurial activity. At the end of the interview the surveyors conducting the interviews were asked about their perception of how easy it was to get hold of the girl they had just surveyed and her willingness to share information.

While a questionnaire is a simple way to reach out to many people in a limited time frame, phone interviews have the additional benefit of not requiring an in-person meeting with each participant. The main advantage of phone interviews is that they enable data to be collected from geographically scattered samples at a lower cost and more quickly than through face-to-face interviews (Purdon & Thomes, 1994).

Disadvantages of phone interviews include the lack of telephone coverage for some participants and the potential for distractions in their environment (Opdenakker, 2006). For example, one might reach a participant when she is at home with children running around and family members doing household chores, factors which might distract her from focusing fully on the phone interview. Another disadvantage of phone interviews is that they, unlike face-to-face interviews, must be kept short, a factor which reduces the potential for in-depth discussion (Sweet, 2002). As our interview guide was designed with the intention of making the interviews short, this did not work as a disadvantage in LTPS.

Another critical problem with phone-based surveys is the challenge of obtaining representative samples when parts of the population do not have a telephone (Purdon & Thomes, 1994). However, in Sub Saharan Africa, phone coverage has increased drastically over the last decade and most people have a telephone at home or a mobile phone (Aker & Mbiti, 2010). To increase the likelihood of reaching the girls after they had finished school, comprehensive contact information was collected on the girls' family and friends – both in the baseline and STS. With up to 18 phone numbers from each girl's network, we expected to be able to reach the girl even if she did not have a phone herself.

Both questionnaires and the interview guide in LTPS were conducted in Swahili. In order to limit translation biases, both questionnaires and the interview guide from LTPS were cross-

⁸ The interview guide can be found in Appendix A.

translated prior to the surveys. The cross-translations involved a translation from the original English versions to Swahili, followed by an independent group of translators converting the Swahili versions back to English. Finally, the two English versions were compared to correct possible errors.

In addition to evidence from surveys, we benefitted from interviews with Jacqueline Mgumia, Dr. Goodluck C. Urassa, and representatives from Femina HIP when analyzing and discussing our findings. Mgumia is a PhD student connected to the University of Dar es Salaam, University of Witwatersrand and the Norwegian School of Economics. She mainly works on issues related to youth, entrepreneurship and poverty. In corporation with the entrepreneurship center at University of Dar es Salaam (UDEC), she has planned and executed an ethnographic study in Tanzania, looking at the relationship between the nature of entrepreneurship training and business outcomes among youths in an urban setting. Dr. Urassa was working as Deputy Director at the University of Dar es Salaam Entrepreneurship Center at the time when the interventions were implemented and contributed his knowledge and experience on youth and entrepreneurship. Femina HIP is the biggest civil media initiative in Tanzania, reaching out to young people and communities across the country. Femina HIP promotes healthy lifestyles, sexual health, HIV/AIDS prevention and gender equality.

5 Analysis

5.1 Data

One of the key outcomes from the Girl Empowerment Project is the behavioral dimension of entrepreneurship, where the girls' current plans of opening a business are measured. We use this outcome, *plans to start a business*, as one of two dependent variables in our analysis.

In order to answer our research question, we are also interested in whether the girls have acted on their plans. By asking the girls whether they have income from work, we received an indication of whether the girls have taken some action to start generating an income. We argue that one must consider both income from self-employment and income from other work, when determining whether the girl has taken action on her plans. As mentioned in section 3.3, self-employment in the informal sector is often a last resort of earning an income, indicating that most people would choose formal employment over starting their own business. Hence, if a girl has income from a formal job, this does not mean that she would not have started a business if she were not given the employment opportunity. This is why we include whether the girl has income from either her own business or from other work when constructing our second dependent variable; *income*.

The girls were asked whether they had plans to start a business both in the STS and LTPS. Questions about income were also asked in both STS and LTPS. Below follows an explanation of the construction of the two dependent variables.

Business plan is a binary variable taking the value 1 if the girl has plans to start a business, and zero otherwise. The question in STS was articulated as follows:⁹

4.3. Have you made plans to start a business recently (in the last few months)?

Income is a binary variable taking the value 1 if the girl has income from work, and zero otherwise. The question from STS was articulated as follows:

4.1. Do you have your income from work?

⁹ The question from LTPS can be found in Appendix A (Q13.1).

In order to capture whether the girl was receiving money from other sources than work, such as pocket money from family members or boyfriends, the question in LTPS was formulated differently than in STS by asking whether the girl has a regular income *in general*.¹⁰ In the follow-up question, the girls were asked to categorize the source of their income as either income from employment, from their own business or from another source. We define income from own business or employment as income from work, which makes the definition of the *income* variable identical in both STS and LTPS. This makes it possible to study the development of incoming-generating activities between the two surveys.

Through questions about constraints to starting a business, we tried to map the challenges the girls are facing. Some of these challenges could be related to the girls' individual characteristics. In particular, human capital aspects seem likely to be crucial determinants of entrepreneurial success.¹¹ For this reason, we include a selection of background variables on family and individual characteristics in order to identify some common features among the girls who had plans to start a business or were receiving an income. The included covariates are *wealthy household*, *cognitive abilities*, *business knowledge*, *risk preferences* and *age*. Below follows a description of how the variables are constructed.

Because poor students perform worse than wealthier students in national tests, family wealth is recognized as a factor that impacts students' academic results (UNICEF; US Aid, 2013). Further, both academic performance and prohibitive school fees are a hindrance for many Tanzanian students to continue their education.¹² As starting a business is often an alternative to pursuing further studies, it is possible that family wealth could be linked to entrepreneurial behavior.

The *family wealth* variable is an index based on the three following questions from the baseline survey:

¹⁰ The question from LTPS can be found in Appendix A (Q10).

¹¹ Personality factors of entrepreneurs have been studied by researchers since the early 1950s. Characteristics like being the oldest child in the family, having self-employed parents, tolerance of risk and creativity have been emphasized by researchers like Timmons (1994) and Bianchi (1993). However, there is weak evidence supporting the research. Although the personality and socio-cultural factors proposed may seem logical, these variables seem to explain only a small part of who will be a successful entrepreneur (Byers, Kist, & Sutton, 1997).

¹² In order to enter higher education the student has to pass the Form IV examinations (CSEE). Even though the student fulfills the academic requirements, the financial aspects withhold some of them as they are not able to pay the required school fees required. See section 3.2.

1.4.1: Do you have a TV at home?

(Yes/No)

1.4.6: How often do you eat meat at home in a normal week?

(Number of days per week)

1.4.10: What is the main source of energy for lighting in the house?

(Tick of one: electricity, kerosene lamp, candles, solar energy, other)

The index is constructed by calculating the average of the standardized variables for these three questions. Further, a binary variable indicates whether the index falls above or below the median.

Cognitive abilities refer to the ability of acquiring knowledge through thought, experience and the senses (Pearsall, 2002). As cognitive abilities can be closely linked to academic performance, it could also give an implication of whether the girl is likely to pass final exams. As previously mentioned, starting a business is considered an alternative career path if not accepted for further studies, which makes this an interesting relationship to investigate further.

The variable is constructed using the same methodology as was used for wealthy household. An index is calculated based on the number of correct answers to questions 2.1.1 – 3 from the baseline survey:

2.1.1 How many zeroes do you have to include if you write “twenty five million” in figures?

2.1.2 The full price of a coat is 250,000 TZS, but in a sale, the price is reduced by 20 %. How much do you have to pay for the coat?

2.1.3 Write the following in order of size, starting with the smallest:

$\frac{2}{3}$	65 %	0.6

A binary variable is then constructed, indicating whether the girl has cognitive abilities above or below the median.

We consider it likely that a certain level of knowledge about business practices is necessary to start a business. To know how profit differs from income, how to treat a customer and how to identify a market is essential, regardless of the size of the business. Yet again, not everyone with business knowledge wants to start a business, thus we are curious whether there could be a relationship between business knowledge and our outcome variables. The *business knowledge* variable is an index that equals the number of correct answers given to three incentivized questions about business in the baseline survey.¹³

For those who have little money in the first place, even starting a small-scale business can involve great risk. Thus we are curious to whether we can observe some correlations between how willing a girl is to take risk and whether she has plans to start a business or have income from own work. *Risk averse* is a binary variable taking the value 1 if the respondent ticks off one of the first two alternatives from question 2.2.3 in the baseline survey:

2.2.3 Imagine that you have 100,000 TZS. You can choose to invest some of the money in a project that doubles your investment if you are lucky, but where you lose all your investment if you are unlucky. There is an equal chance of being lucky and unlucky. How much of the 100,000 TZS would you invest? Tick off one.

I would keep 100,000 and invest 0	I would keep 75,000 and invest 25,000	I would keep 25,000 and invest 75,000	I would keep 0 and invest 100,000

We know that the girls' *age* vary within each grade, and we consider it an interesting variable for a couple of reasons. Firstly, it is possible that older girls are more mature, and thus better suited to start a business than their younger classmates. At the same time we know there might be underlying reasons for why some of the girls are older, such as retention and poor family situations causing delays in girls' participation in school. The *age* variable is constructed as a binary variable, indicating whether the girl is older or younger than the median of 17.

¹³ The questions can be found in Appendix D.

5.2 Econometric specifications

In the subsequent regressions on business plans and income we are using ordinary least square (OLS). OLS is widely used to model the dependent variable Y as a linear function of the independent variables X . However, there are some shortcomings to the OLS model when estimating outcomes on a binary variable. When using a binary variable as the outcome variable, the regression line will end up going below 0 and above 1 - not fitting the data very well. To confirm whether the results from OLS estimations are robust, we also run the same regressions using a probit model. Probit regression, also called a probit model, is used to model dichotomous or binary outcome variables. In the probit model, the inverse standard normal distribution of the probability is modeled as a linear combination of the predictors. A link function $F(Y)$ is needed for this transformation, which uses the cumulative normal distribution to produce probabilities between 0 and 1 for any given value. This link function is known as the Probit link (Schafer & Ramsey, 2002). The estimated curve is an S-shaped cumulative normal distribution that fits the data much better than a linear estimation, and always lies between 0 and 1.

In developing countries like Tanzania, there are differences between rural and urban areas when it comes to poverty, education, birth rates and economic activity.¹⁴ With lower academic performances in remote areas and less exposure to economic activity, there could be differences in how the girls respond to the various training programs in remote and non-remote areas. We consider the possibility that a sample consisting of only remote schools will better represent the four schools from LTPS, as they are all remote.¹⁵ In our analysis from STS we thus run separate regressions for the full sample, for only remote schools and the four schools from LTPS.

We first run a regression on the outcome of interest (the likelihood of having plans to start a business and having income from work) by treatment status. We include a binary variable for each of the three treatment groups; health, business and the combined treatment. The health training is not expected to have an impact on neither business plans nor income, given its

¹⁴ This is discussed in Chapter 3.

¹⁵ *Remoteness* is a binary variable taking the value 1 if the time it takes to reach the school by car from the local district headquarters is above the median of the distribution. Thus, the variable identifies the 50 percent most remote schools.

purpose of empowering girls on gender issues and sexuality. However, in order to state the relative importance of the treatments, we include it in the regression:

$$(1) Y_i = \alpha + \beta_1 Health_i + \beta_2 Business_i + \beta_3 Combined_i + \varepsilon_i$$

The error terms are clustered at the school level.

Because there may be omitted variables that are correlated with our outcome variables¹⁶, we also estimate regression (1) with a set of covariates X presented in section 5.3:

$$(2) Y_i = \alpha + \beta_1 Health_i + \beta_2 Business_i + \beta_3 Combined_i + \beta_4 X_i + \varepsilon_i$$

¹⁶ Omitted variable bias occurs when a variable that has been omitted from the analysis is correlated with X , and the omitted variable is a determinant of Y . It constitutes a violation of the assumptions behind OLS and makes the OLS estimator inconsistent (Stock & Watson, 2012).

6 Call history from the long-term pilot survey

This chapter describes the process of reaching the girls in the LTPS, which was one of the two purposes of the survey. The intention of analyzing the data from the call history is to try to map how easily the girls can be reached and how many one can expect to reach in the long-term survey. When analyzing the call history, we first present the results from the full sample. We also look at differences between those who were offered one of the three training programs (the health, business and health & business treated schools) and those who were left untreated (the control school). By making this distinction, we are hoping to uncover potential differences in how easy it is to reach the girls and how willing they are to share information; could girls from the control group be harder to reach or less inclined to share information than those who were given training from the project? Table 2: Call history shows a summary of findings from the call history data. Further we will discuss findings presented in the table, and supplement them with other qualitative information we have from LTPS.

How many did we reach?

There were 209 girls from all four schools in total. We managed to reach 179 out of the 209 girls, which is equivalent to 85 percent. There is no significant difference between trained and non-trained in this regard; we reached 86 percent of the girls who received a type of training and 83 per cent from the control school.

How much time did it take to reach one girl?

To make a rational estimate of how much time it took to reach one girl, the average is calculated based on recorded working hours for the surveyors in Dar es Salaam. Although some of this time was used for various discussions and problem solving, we would expect interruptions like these to be present also for another phone interview session, thus we expect the estimate to be realistic and applicable for the long-term survey. To reach and execute an interview for *one girl* took on average *one hour and five minutes*.

Table 2: Call history

Treatment:	All	Treated	Control
How many reached?	85 %	86 %	83 %
Time to reach one girl (hours)	1.08		
Time per interview (minutes)	12.3		
Calls per girl	6.3	6.5	5.4
<i>Reached through:</i>			
· Family	43 %	45 %	38 %
· Self	50 %	49 %	54 %
· Friends and other	6 %	6 %	8 %
Easy to reach (1-5)	3.7	3.5	4.5
Willingness to share (1-5)	4.8	4.8	5.0
Observations targeted	209	162	47
Observations reached	179	140	39

Note: The table shows a summary of findings from the call history. For some findings we present findings separately for girls that have received training (health, business or combined treatment) and those who have not (control group). The time it takes to reach one girl is an average calculation based on recorded working hours. The time per interview is calculated from automatic time registrations for when the first and last question was answered. Calls per girl show the average of phone calls we had to make before reaching or not reaching the girl. The surveyors' perception of how easy the girls were reached and how willing they were to share information is shown in the two bottom rows, where 1 represents "very difficult" or "very unwilling" and 5 represents "very easy" or "very willing".

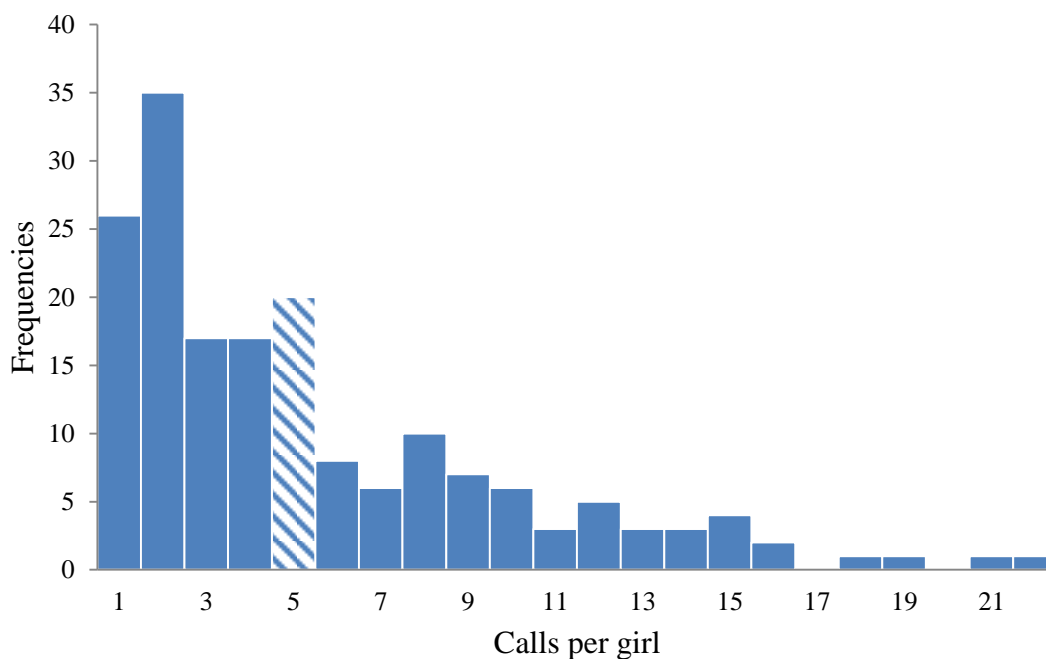
Time per interview

The questionnaire had automatic time-variables embedded which recorded the times for when all questions were entered into the excel sheet. By subtracting the times for when the last question was entered into the excel sheet from when the first question was entered, we were able to calculate the average time spent on an interview. From Table 2: Call history we can see that an interview took 12 minutes on average.

How many calls per girl?

Our data show that we made on average six calls to reach each girl. Naturally there were some girls requiring more calls than others. We have illustrated this in Figure 2 which shows the frequencies for number of calls we had to make to reach each girl or defining the girl as unreachable. In 26 cases, one call was sufficient to reach the girl, while one girl required 24 calls before the surveyor dismissed her as unreachable. The most frequent number of calls was two, and the median of five calls is highlighted with a striped bar in Figure 2.

Figure 2: Calls per girl



Note: The histogram shows the frequencies for each number of calls the surveyors made to reach a girl or defining her as unreachable. The median is outlined with the striped bar.

Who did we reach the girls through?

Half of the girls that we reached were reached on their personal phone numbers. Various family members like parents, siblings, aunts and grandmothers helped us get in contact with 43 percent of the girls, while 6 percent were reached through friends, neighbors and teachers. The distribution between personal numbers, family members and friends are not much different when comparing the girls who received training to those who were left untreated. From the treated sample, 49 percent were reached on their personal phone numbers, 45 percent were reached through family members, and 6 percent through friends and other. When trying to get in touch with girls from the control school, we reached 54 percent on their personal phone numbers, 38 per cent through family members, and 8 per cent through friends and other.

It is important to mention that the surveyors on several occasions received incoming calls, which resulted in an interview. The girls could be returning a missed call on their phone, but could also be calling as a result of family, friends or others bringing them the message that we were trying to get hold of them. Unfortunately, these calls were not registered in the call history as it was not designed appropriately to record incoming calls.

Also worth noting is that teachers helped the surveyors more than the data show. On several occasions the surveyors contacted the teacher and asked for their help to track down the girls. The teachers provided us with information on where the girls were staying and both personal phone numbers and phone numbers for the girls' relatives. They also sent other students to deliver our message to the girls that we were trying to reach them.

Were they easy to reach and willing to share?

After finishing each interview, the surveyor was asked about his or her perception of how easy it was to reach the girl and how willing the girl was to share information. These are found as Q17 and Q18 in the Questionnaire. The surveyors did so by choosing a number on a scale from 1 to 5, where 1 represents "very difficult" or "very unwilling" and 5 represents "very easy" or "very willing". From these ratings we were able to see the surveyors overall perception of the process.

From Table 2 we see that the girls were on average *easy* to reach (4 indicated "easy"). It seems like the girls from the control school were even easier to reach, with an average ranking

of 4.5, compared to the treated schools with an average ranking of 3.5. While the treated schools were leaning towards *neutral*, the control school was leaning towards *very easy*.

With a score of 4.8 on a scale from 1 to 5, most of the girls were *very willing* to share information. This also counts for the schools that received training. From the control school, all girls were ranked as 5, i.e. they were all *very willing* to share information.

7 Findings

The primary purpose of the pilot was to gather updated and more in-depth data on a subsample of the girls from the Girl Empowerment Project, in order to know more about their daily lives and activities: how are they spending their days and making a living? Have they attempted to start a business or do they have plans of doing so? Are there constraints holding them back – and if so; what are they? The first section in this chapter presents findings from phone interviews in LTPS. In section two and three we present results from regressions on whether the girls have plans to start a business and income from work. We have estimated the regressions on data from both STS and LTPS, and we draw links between the two. We discuss the findings in light of insights provided to us by the literature review and institutional context/framework, along with interviews which we conducted with several experts in the field while visiting Tanzania.

7.1 Findings from phone interviews in LTPS

Answers to questions on the girls' family status, their current residence, and their results from the Form IV examination is presented in Table 3. Below follows a short explanation of the findings.

Family

Most of the girls are not yet married. However, while only 3 percent reported being married, 17 percent of the girls consider themselves to be in a permanent relationship.¹⁷ One girl has children, and three girls reported being pregnant. To conclude, it appears that only a few of the girls have decided to start a family at this point.¹⁸

¹⁷ Following the recommendation from Femina HIP, we included question 9.1.2 "Are you in a permanent relationship?" to those who reported not being married in question 9.1.1. As the question was added later, the question was only asked to 120 of the girls.

¹⁸ Table C1 in Appendix C, shows that there are no significant treatment effects on marriage and permanent relationships from the various training interventions. Regardless, any potential treatment effects would have been hard to identify with such few observations.

Moving

Since the STS, 31 percent of the girls have moved. In later sections, we will elaborate on reasons that were given regarding *why* they have moved. According to the girls

Table 3: Family status and education

Family status/education	All
Married	3 %
Permanent relationship	17 %*
Children	1 %
Pregnant	2 %
Moved	31 %
Passed Form IV	14 %
Observations	179

Note: The table shows the percentage of girls for which the various statements are true. Only 16 girls (9 %) do not know their results from the Form IV examinations. *This number is based on 120 observations because the question was included after the phone interview sessions had started.

there are a variety of reasons for moving; some have moved to pursue further studies, others have moved to stay with family members and look for opportunities in business, and some have moved back home to their parent's homes after having stayed with relatives during the school semesters. Girls who have moved to grandparent's homes have done so to provide care for them and help out with household activities, while girls who stayed with sisters or aunts have done so in search of business or educational opportunities.

Form IV Examinations

All 179 girls sat for the Form IV Examinations (CSEE). However, as registration fees have not been paid, nine percent of these girls do not know their results because government holds back exam results until the registration fee and additional late fees are paid.¹⁹ Out of those who *did* know their results, 14 percent passed the exam. Compared to the 2013 national average of 58 percent,²⁰ this is incredibly low. However, as the national average does not distinguish between results from public and private schools, this is most likely not a valid comparison.

Out of the 23 girls who passed the Form IV examinations, only two have moved on to college. They are both in nursing colleges, while two other girls are waiting to go to Tourism College and Teaching College. While waiting, they are doing daily work around the house or helping out with family business. None of the girls have continued their secondary education by starting Form V.

Daily activities

To create a better picture of how the girls spend their days, we included an open-ended question regarding their daily activities. By categorizing the answers into groups of similar answers, we see that about half of the girls report doing household chores as their main daily activity.

Of the remaining girls, quite a few reported other main activities *in combination* with household chores. These are activities such as vocational courses, helping out in the family's business, and farming. There are also girls who work outside their home, generating an income. By asking the girls whether they have a regular income – and if so, the size and source of it – we were able to see what kind of income-generating activities the girls mostly engage in and which activity generates more income.

Income

In Table 4, we present the share of girls who reported having a regular income, and indicate whether this stems from employment, their own business, or another source. It also shows the average income overall, and the average income from the three different categories. The table

¹⁹ See Chapter 3 for more details registration fees.

²⁰ See section 3.2

shows that 23 percent of the girls reported having a source of regular income. While 6 percent reported that their regular income come from a formal job, 17 percent report that they receive income from their own business. Thus, running their own business seems to be the way most girls generate an income. As we were expecting to find some girls receiving regular income

Table 4: Regular income and farming

		Obs.
Regular income, total	23 %	179
· Regular income from employment	6 %	179
· Regular income from own business	17 %	179
· Regular income from other sources	1 %	179
Average income, total	19 000	42
· Average income from employment	15 900	42
· Average income from business	20 800	42
· Average income from other sources	1 500	42
Farming	52 %	120
· Farming on own or rented land (“own business”)	20 %	120
· Farming on parents’ land	32 %	120

Note: The top figure shows the percentage of girls who report having a source of regular income. “Regular income, total” is the sum of the three categories employment, own business and other. The next four figures show the average income among those who report having regular income; the total average and the average within the three categories. All amounts are in Tanzanian Shillings (TZS) and are rounded to the closest 100. Finally we show outcomes from questions on farming. The questions were asked to 120 of the girls.

from other sources, including financial contributions from family members or boyfriends, we added a third open category; *'other sources'*. Only one girl reported receiving regular money contributions from a family member, which appears, in the table, as the 1 percent with regular income from other sources.

For those with a source of regular income, the average income is 19,000 Tanzanian shillings (TZS) per week. For the girls with regular income from employment, the average income is 15,900 TZS per week. The average income is highest among girls who have their own business, with an average weekly income of 20,800 TZS. According to Maalim, an average meal consisting of rice and meat cost roughly 2,000 TZS in remote areas in Morogoro.

Among the girls who have income from their own business, businesses include anything from selling teacakes or charcoal, to sewing clothes and braiding hair. The highest regular income from a business is 70,000 TZS per week and is received by a girl selling charcoal. This is a high income compared to other charcoal selling girls who have an average income of 14,000 TZS. It is also 30,000 TZS higher than the second highest income, which was reported by a girl from the control school who is also selling charcoal. This second girl started her own business transporting charcoal from Morogoro to Dar es Salaam to sell it, with an average income of 40,000 TZS per week. However, she is temporarily out of business because of a broken bridge along her travel route. The least profitable business belongs to a girl braiding hair, which generates an average income of 2,500 TZS per week.

Farming

Questions regarding farming were included in the questionnaire after the phone interview sessions had started, which resulted in fewer observations.²¹ The 120 answers we received are still valuable, as they provide us with information about the extent to which the girls perceive farming as a business or not. These findings may have further implications for the design of future training programs.

The results of the questions regarding farming are presented in Table 4, and shows that 52 percent of the girls asked are involved in farming activities. We consider farming a business if the girl rents or owns the land – regardless of how she ended up owning it. Thus, out of the

²¹ As the health and business treated schools were on the top of the surveyors calling lists, most of the missing observations come from these two schools.

girls involved in farming activities, 20 percent farm as business, while the remaining 32 percent are helping out on their parents' land.

Business plans and implementation of business plans

Table 5 shows our findings from LTPS on the girls' business plans and involvement in business. The question "do you have plans to start a business?" was also asked in STS. At the time of the STS, 36 percent of the girls from the STS sample said that they had business plans. In the LTPS, 40 percent of the girls report having plans to start a business.

Table 5: Business plans STS and LTPS

	Business plan LTPS: No	Business plan LTPS: Yes	Total
Business plan STS: No	43 %	22 %	64 %
Business plan STS: Yes	17 %	18 %	36 %
Total	60 %	40 %	100 %

Note: The table shows the percentage of girls with plans to start a business and no plans to start a business in STS and LTPS. The percentages are based on a sample of 166 girls.

Table 6: Business plans STS and income LTPS

	Income LTPS: No	Income LTPS: Yes	Total
Business plan STS: No	83 %	17 %	100 %
Business plan STS: Yes	75 %	25 %	100 %

Note: The table shows the percentage of girls who did/did not have income in LTPS among those who had/did not have business plans in the STS. “Income” is income from either own business or employment.

This indicates that the girls still have business plans – indeed, a few more have plans now than directly after the training. However, these two figures do not draw the whole picture of the continuity of business plans, nor do they indicate whether it is the same girls who report having business plans in LTPS as in STS. From the table we also observe that 18 percent had business plans in both STS and LTPS, i.e. about half of the girls with business plans directly after the interventions still have plans to start a business, and about half of them have let go of their plans. At this point, we cannot state anything about the reasons for the observed changes in business plans.

In Table 6, we investigate whether having a business plan at the time of the STS could have made the girls more likely to report having income from work in the LTPS. While 17 percent of those with no business plan in STS reported having an income from work in LTPS, 25 percent of those *with* business plans in STS reported having an income in LTPS. Hence, the percentage who reported having an income in LTPS is higher among those who had business plans in STS, than among those who did not have plans in STS. We run a two-sided t-test to determine whether the difference in means of income is different from 0. We find that there is no significant difference in the means of income for girls with business plans in STS ($M = 0.25$, $SD = 0.44$) and those with no business plans in STS ($M = 0.17$, $SD = 0.38$) conditions; $t(190) = -1.28$, $p = 0.20$. These results suggest that having business plans in STS did not make girls more likely to have income from work in LTPS.

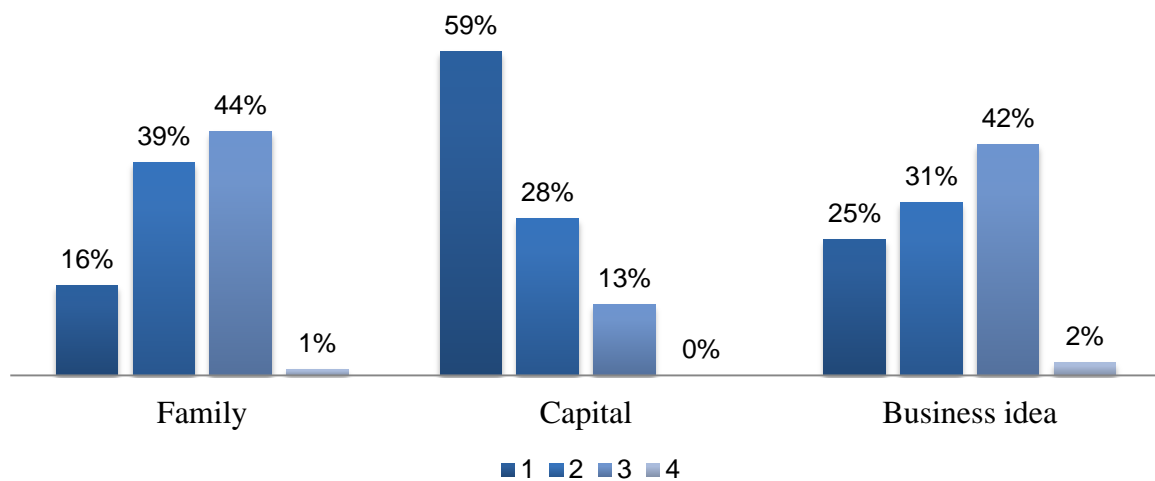
Constraints

What happened to the 17 percent of girls who no longer have a business plan? As mentioned in section 6.1.1, very few have continued their education after graduating from Form IV. In section 6.1.2, we show how some of the girls have income from employment and their own

business, and some girls expect income from future harvests. Still, 36 girls out of the 59 with business plans in STS have given up on their plan.²²

To gain a better understanding of what the girls perceive as their biggest constraints to starting or running a business, the girls were asked to rank three constraints: family obligations, start-up capital and business idea. They were also given the option to specify an additional constraint and assign a rank to it. Constraints were ranked from 1 to 4, where 1 represented

Figure 3: Constraints to starting a business



Note: The histogram shows how girls rank three various constraints to starting a business; family obligations, start-up capital, and business idea. Various other constraints are not included in the histogram. Constraints were ranked from 1 to 4, where 1 represents the most important constraint and 4 represent the least important constraint.

the most important constraint and 4 represented the least important constraint. In addition to ranking the constraints they were also asked to specify the reasoning for each rank assigned.

Figure 3 shows how the full sample ranks the constraints *on average*. From the histogram we see that most girls consider capital as the most important constraint. What constraint is considered most important between family and business is not as clear, but more girls consider business ideas as their number one constraint, than those who consider family to be their primary constraint. The constraints reported under “other” were mostly ranked as the

²² There are 16 girls reporting income from either own business or employment, and 7 girls are farming on own or rented land. The number of girls farming on own or rented land could be higher as only 120 girls were asked about farming.

least important constraint, if reported. As only a few reported another constraint, it is not included in the histogram.

How can constraints stop implementation of business plans?

Constraints to running a business that are linked to family obligations include the various family responsibilities women have in the household, such as taking care of children and cooking. With several obligations at home, many of the girls do not have the time to run a business – which often requires working from early morning until late evening. Some of the girls point out that the family responsibilities may lead to a collapse of the business if the capital is used to cover the family's needs. One girl states: “... *the typical situation is when you have opened up a business, but all family members are depending on that business to have their necessary needs – like food. So the business will probably close*”.

The girls rank capital as the most important constraint to start or run a business. They argue that one cannot start a business without capital, and that one needs capital to even plan to start a business. Others explained that without sufficient capital, the business will not grow, and this is often why businesses collapse. While most girls agree on the importance of capital to start a business, a few argued that it is possible to start a business even with low capital: “*one can start small with small capital*”.

The importance of a business idea as a constraint to starting and running a business was explained in various ways. Some girls stated that many business ideas have already been implemented, which makes it difficult to come up with new ideas. Others made the point that business ideas are determined by the size of capital, making constraints from lack of capital visible through a business idea with little potential. A few of the girls consider business plans a smaller constraint and stated that it is easy to come up with business ideas as everything is doable. One girl mentioned that there are many ideas, but without capital and with family obligations, the idea is nothing.

“Other” constraints mentioned included lack of business skills, lack of access to business advice, and issues related to selling products on credit when customers do not pay you back.

In an open-ended question where the girls were asked about their business plan and whether it has changed since the STS, we received further explanations on how various constraints prevents them from implementing their plans. The main reasons why the girls have changed

their original business plans are lack of capital and missing skills. These obstacles have made some girls give up on their plans, while others have simply postponed their plans, while saving money or acquiring the skills needed. One girl wanted to open a clothing shop, but with limited capital, she has started a business by selling vegetables and is now saving some of the profit to establish her planned business. With financial help from her mother she now has got 80,000 TZS in savings and is optimistic about opening the clothing shop in the near future. Another girl explains that her business plan changed since the STS when she learned that she would get more customers if she opened a home appliance shop, rather than selling female clothing. Fear of customers not paying or delaying the payment led another girl to let go of her plan of opening a hairdressing saloon.

7.2 Treatment effects on having plans to start a business

In the following sections we present regression results from STS on plans to start a business and income. Table 7 reports the results from regression (1) and (2) presented in section 5.2 on plans to start a business. We run the same regressions on the full sample, on a subsample consisting only of girls from remote schools, and on the four school sample. Columns (1), (3) and (5) present results from regressions on average treatment effects from the health, business and combined treatment. In the regressions in columns (2), (4) and (6) we have included covariates to control for various background characteristics of the girls.

All treatment coefficients for the business treatment are positive, indicating that girls were more likely to have business plans in STS if they had received the business training. We observe a slightly greater treatment effect from the combined treatment than from the business treatment. This implies that girls who received the combined treatment were even more likely to have business plans in STS than those who received only the business training. Given that the girl received either business training or the combined treatment, there are only marginal differences in the probability of having business plans in STS when comparing the results from different samples. However, we do see a slightly higher effect from the combined treatment in the regression in column (5). If a girl from the four school sample received the combined treatment, the predicted probability of her having a business plan was 0.744, while the same probability for the STS remote sample was 0.583, as shown in the regression in column (3). All treatment coefficients for the business and the combined treatment are significant at a 1 percent level.

The treatment coefficients for the health treatment are positive in all regressions, but only significant in the remote sample. However, from the regression in column (3) we see that girls from remote schools who received health treatment were only 6 percentage points more likely to have business plans than girls who did not receive any training. Although significant, the effects on business plans from health treatment are also relatively small among girls from remote schools.

When looking at the results from the regressions with control variables, we notice that the coefficient for wealthy household is positive in all regressions. The coefficient from the four school sample implies that girls were about 13 percentage points more likely to have business plans in the STS if they come from a wealthier family. However, the coefficient

Table 7: Business plans (STS)

	Full sample		Remote sample		Four school sample	
	(1)	(2)	(3)	(4)	(5)	(6)
Health	0.009 (0.030)	0.018 (0.028)	0.061** (0.028)	0.059** (0.024)	0.129 (0.085)	0.148 (0.093)
Business	0.361*** (0.041)	0.373*** (0.039)	0.369*** (0.027)	0.374*** (0.028)	0.472*** (0.092)	0.552*** (0.094)
Health & business	0.399*** (0.045)	0.427*** (0.045)	0.487*** (0.055)	0.506*** (0.048)	0.716*** (0.089)	0.701*** (0.097)
Wealthy household		0.027 (0.022)		0.039 (0.027)		0.128 (0.082)
Cognitive abilities		0.052*** (0.018)		0.036* (0.021)		0.099 (0.067)
Business knowledge		0.041*** (0.011)		0.037** (0.015)		-0.029 (0.046)
Age > 17		0.006 (0.017)		0.002 (0.023)		-0.107 (0.066)
Risk averse		0.015 (0.017)		0.017 (0.021)		0.038 (0.069)
Constant	0.163*** (0.022)	0.040 (0.030)	0.096*** (0.016)	0.000 (0.033)	0.028 (0.065)	-0.011 (0.117)
R^2	0.153	0.181	0.182	0.206	0.345	0.394
t_HB	0.370	0.391	0.431	0.433	0.601	0.700
se_HB	0.059	0.055	0.045	0.042	0.156	0.163
Observations	3305	2849	1691	1459	166	138

Note: The table reports estimates from OLS regressions for the average treatment effects on plans to start a business, where the dependent variable is an indicator variable taking the value 1 if the girl had business plans in the short-term survey. Columns (1) and (2) report results for the full sample. Columns (3) and (4) report outcomes for only remote schools. Columns (5) and (6) report results for the four school sample. "Health", "Business" and "Health & business" are treatment variables taking the value 1 if the girl was assigned to this treatment. t_HB and se_HB shows the point estimates and their belonging standard errors for a linear combination of the "Health" and "Business" coefficients. Differences in sample size when covariates are included reflect missing information from some of the participants. Standard errors in parentheses, clustered at school level for the remote and full sample. Standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

is not significant. The coefficients for both cognitive abilities and business knowledge are found to be positive in all regressions, and significant in the regressions for the remote and the full sample. The estimated coefficients are greater and more precise in the full sample, which shows that that if the girl had a score above the median in the cognitive abilities index or from the business knowledge questions, she would be about 5 and 4 percentage points more likely to have business plans in STS, respectively. As we can think of explanations for both a positive and negative sign of these coefficients, it is hard to argue for their reasonability. If being smarter makes you more capable of coming up with a business plan, a positive relationship could make sense. If the girl is smart she could also be more inclined to prioritize studying over starting a business, which speaks for a negative coefficient.

The results on the treatment effects in STS are in line with our expectations, as we were expecting to see more girls with plans to start a business among those who had received business training, than among those who only received the health treatment. As a robustness check we perform a probit analysis with the same variables as in the OLS regressions. The marginal effects from the probit estimations show a similar pattern to the one we find in the OLS estimates.²³

Table 8 reports the results on having a business plan in LTPS. Column (1) and (2) present the results from OLS regressions with only treatment variables, and with treatment variables and covariates. Column (3) and (4) report the marginal effects from probit estimations.

The regressions reported in column (1) and (2) in Table 8 show that both the business training and the combined training seem to have had positive long-term treatment effects on having plans to start a business. From the regression in column (1) we find that the probability for a girl from the control group to have plans to start a business is 0.154. The probability for a girl who has been offered business training is 0.583, which is slightly lower than the probability for a girl who has been offered the combined treatment, which is 0.622. Both the business and combined treatment coefficients are significant at a 1 percent level. The findings from LTPS shown in column (1) and (2) are also robust when comparing them with the marginal effects from the probit estimations presented in column (3) and (4). The explanation for this strong treatment effect from the business training and the combined treatment can be found in the literature, which argues that human capital interventions, like business training, can speed up

²³ Marginal effects from probit estimations are reported in Appendix B: Probit estimations.

Table 8: Business plans (LTPS)

	OLS (1)	OLS (2)	Probit (3)	Probit (4)
Health (d)	0.100 (0.093)	0.036 (0.098)	0.138 (0.116)	0.056 (0.127)
Business (d)	0.429*** (0.104)	0.428*** (0.105)	0.461*** (0.105)	0.485*** (0.116)
Health & business (d)	0.468*** (0.099)	0.460*** (0.107)	0.494*** (0.098)	0.516*** (0.115)
Wealthy household (d)		-0.204** (0.086)		-0.225*** (0.083)
Cognitive abilities (d)		0.066 (0.072)		0.071 (0.085)
Business knowledge		-0.001 (0.050)		-0.005 (0.059)
Age > 17 (d)		0.111 (0.072)		0.135 (0.083)
Risk averse (d)		0.080 (0.074)		0.097 (0.088)
Constant	0.154** (0.072)	0.070 (0.125)		
R^2	0.165	0.243		
t_HB	0.530	0.464	1.589	1.461
se_HB	0.173	0.177	0.560	0.619
Observations	179	151	179	151

Note: The table reports average treatment effects on plans to start a business, where the dependent variable is an indicator variable taking the value 1 if the girl had business plans in LTPS. Columns (1) and (2) report outcomes from OLS regressions and columns (3) and (4) reports marginal effects from probit estimations. "Health", "Business" and "Health & business" are treatment variables taking the value 1 if the girl was assigned to this treatment. t_HB and se_HB shows the point estimates and their belonging standard errors for a linear combination of the "Health" and "Business" coefficients. Differences in sample size when covariates are included reflect missing information from some of the participants. Marginal effects; Standard errors in parentheses, (d) for discrete change of dummy variable from 0 to 1, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

the entry for women interested in starting a business. Making plans to start a new business is the first step to starting a business.

The observed treatment effects from the business and combined treatment in LTPS are line with the findings and expectations from STS, although both treatment effects are slightly lower in LTPS. The treatment effect from the combined treatment is still higher than the effect from business training, but the difference is less pronounced in LTPS than in STS. The point estimates from a linear combination of the health and business treatment coefficients supports this: in STS, the coefficient from the combined treatment is consequently higher than the estimates for the health and business coefficients combined, while in LTPS, they are consequently lower. In STS, there may have been an “additional” effect from the combined treatment, but we do not see this in LTPS. The standard errors point towards the coefficients not being statistically different from each other, though we do not test whether this is the case.

Maalim (2014) compares quantitative and qualitative data on business plans from STS for the four school sample, and identifies a mismatch considering the treatment effects. Contrary to the findings from the data analysis, the girls’ personal essays reveal that business treatment has had a stronger impact on business plans. Although many of the combined treated girls claim to have business plans, the qualitative analysis reveals that these plans are not very thought through. The girls from this group only briefly mention that they want to start a business, compared to the business treated girls who to a larger extent write specific plans in their essays, which reflects that they have thought deeply about starting a business as a part of their future plans. This indicates that the combined treated girls may be more likely to drop their plans, while the business treated girls to a larger extent may keep their business plans. This predicted trend corresponds to what we observed in the LTPS.

The health treatment has little effect on having plans to start a business. This finding is not surprising, as the purpose of the treatment was to provide information about issues related health and gender equality. However, there are significant treatment effects from health training in the remote sample from STS. A possible explanation for this is that it could simply be an effect of being *treated* – we observe a spillover effect. If the health training has succeeded in empowering the girls, this may have led to more girls making business plans. In other words, it is not necessarily a carefully thought through plan of starting a business, but rather a response to an easy tick-off question that reflects a willingness to do something with their life. This is in line with Maalim’s (2014) analysis of the girls’ personal essays, which

states that health treated girls to some extent mention business plans in their essays, but the plans are very general and do not state any specific business plans, or how they plan to put them into life.

Regarding the control variables, there is one coefficient that stands out, namely the negative coefficient for wealthy household. The coefficient implies that if a girl comes from a wealthy household, she is 20 percentage points less likely to have business plans in LTPS. The same coefficient from the STS told us a whole other story, namely if a girl comes from a wealthy household, she is 13 percentage points more likely to have business plans. Although the coefficient was found to be insignificant in the STS, this implies that there must have been girls from wealthier households who have let go of their plans and girls from less wealthy households who have made new plans since the STS.

An explanation for the link between family wealth and having plans to start a business could be that wealthier families have more money to finance further education, which makes the girls less inclined to start a business. However, from our data we observe that the wealthier girls do not continue their studies. Hence, there must be other underlying factors explaining the observed replacement. One explanation could be that girls from wealthier families are being cared for and supported by their family, and thus do not have to start a subsistence enterprise in order to feed themselves and their family. These girls may be happy staying at home, and have no plans to start a business. Poor girls, on the other hand, might experience the opposite. After finishing school, self-employment might stand out as their only alternative, with no further education or employment available. Unlike girls from wealthier families, girls from poorer households might be expected to contribute to the welfare of the household. As a result, poorer girls may to a greater extent realize the need to start a business after leaving school. In Chapter 3 it is stated that subsistence enterprises are mainly run by entrepreneurs who are poor, low educated and low skilled. This underlines our findings of small-scale entrepreneurs mainly coming from poor households. As indicated in Chapter 3, entrepreneurs of small-scale enterprises mostly choose this career path as a last resort for earning a living, and household wealth impacts how urgent it is to get an income.

In discussions with Jacqueline Mgumia, a researcher from the University of Dar es Salaam, on youth and entrepreneurship in Tanzania, the importance of the circulation of money in Tanzanian families was emphasized: *“With no insurance and no guarantee of income every day, money shifts through the family structure because there are not enough resources. There*

is a flow of money through these relationships, and one has to contribute if one believes in this structure.” In other words, it is essential to be aware of the Tanzanian and African culture of community and sharing in family situations with lack of money. Although the community itself is poor, money will be handed to the ones that are truly in need, with an unwritten expectation that the person somehow contributes back to the community when the resources are available. This way no one is left alone, but it also implies great responsibility to share whatever you have available. Hence, a young and healthy girl who has just finished Form IV may feel the push to get an income if the family obligations and household chores are not taking up all her time.

7.3 Treatment effects on having income from work

Table 9 reports the results from OLS regressions for the average treatment effects on having income from work. We run the same regressions for the full sample, for a subsample consisting only of girls from remote schools, and for the four school sample. Columns (1), (3) and (5) present results from regressions on average treatment effects from the health, business and combined treatment. In the regressions in columns (2), (4) and (6) we have included covariates to control for various background characteristics.

The results imply that there were not many girls who had income in STS in general, as the coefficients are low overall. From the regression in column (1) we see that only 4 percent of girls in the control group report having an income. However, as there are very few adolescents in Tanzania that engage in income-generating activities while still in school, the findings are not surprising.²⁴

Although treatment effects are small, the business training and the combined treatment both seem to have had a significant positive effect on income for girls from the full and remote sample. Looking at the regression in column (3), we see that girls who received either business training or the combined treatment were about 3 percentage points more likely to have income from work than those from the control group. The business and combined treatment coefficients are all significant at a 5 percent level, except for the business treatment coefficient in the regression in column (4), which is also significant at a 1 percent level.

²⁴ According to Femina HIP, few secondary school students are involved in income generating activities, and most are dependent on support from their family them while they are still in school.

Table 9: Income (STS)

	Full sample		Remote sample		Four school sample	
	(1)	(2)	(3)	(4)	(5)	(6)
Health	-0.004 (0.009)	-0.005 (0.009)	0.009 (0.012)	0.015 (0.011)	0.020 (0.049)	0.033 (0.050)
Business	0.032** (0.013)	0.029** (0.013)	0.029** (0.011)	0.031*** (0.010)	0.083 (0.053)	0.090* (0.051)
Health & business	0.031** (0.015)	0.036** (0.018)	0.033** (0.016)	0.038** (0.016)	0.119** (0.051)	0.058 (0.053)
Wealthy household		0.011 (0.010)		0.001 (0.011)		0.066 (0.044)
Cognitive abilities		0.004 (0.009)		-0.007 (0.013)		0.011 (0.036)
Business knowledge		0.003 (0.005)		-0.000 (0.006)		-0.004 (0.025)
Age > 17		0.008 (0.010)		0.005 (0.010)		-0.003 (0.036)
Risk averse		-0.000 (0.009)		-0.001 (0.013)		-0.010 (0.037)
Constant	0.040*** (0.006)	0.022* (0.012)	0.027*** (0.005)	0.024* (0.014)	-0.000 (0.037)	-0.010 (0.063)
R^2	0.005	0.007	0.004	0.006	0.044	0.050
t_HB	0.028	0.025	0.038	0.046	0.103	0.123
se_HB	0.018	0.018	0.018	0.016	0.089	0.088
Observations	3304	2849	1690	1459	165	138

Note: The table reports estimates from OLS regressions for the average treatment effects on having income from work, where the dependent variable is an indicator variable taking the value 1 if the girl reported having income from work in the short-term survey. Columns (1) and (2) report results for the full sample. Columns (3) and (4) report outcomes for only remote schools. Columns (5) and (6) report results for the four school sample. "Health", "Business" and "Health & business" are treatment variables taking the value 1 if the girl was assigned to this treatment. t_HB and se_HB shows the point estimates and their belonging standard errors for a linear combination of the "Health" and "Business" coefficients. Differences in sample size when covariates are included reflect missing information from some of the participants. Standard errors in parentheses, clustered at school level for the remote and full sample. Standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The treatment effects in the four school sample are relatively higher than for the full and remote sample, which indicates that differences in income across schools in the four school sample are higher than the observed differences among treatment groups in the full sample. The treatment coefficients from the regression in column (5) in Table 9 implies that girls from the four school sample who received the business training were about 8 percentage points more likely to have income in STS than those who did not receive any training. If a girl had received the combined treatment, the probability of her having an income in STS increased by another 4 percentage points, compared to if she only received the business training. Although it is positive and of a considerable size, the treatment coefficient for business is not significant, which is probably due to small sample size. The coefficient for the combined treatment is significant, but becomes insignificant when controlling for background characteristics. Neither of the coefficients for business and the combined treatment in the four school sample is significant when doing a robustness check with marginal effects from probit estimations.²⁵ The treatment coefficients' substantial sizes do however suggest that there is a difference in income across treatment groups in the four school sample. If there truly is a difference between the four schools, we are expecting to see them more clearly when we do the same analysis on data from LTPS, as we expect more girls to be engaging in income-generating activities now that they are out of school.

We find no effects from the health treatment and detect no consistent relationships between the various background characteristics and whether the girls have income from work in STS. The findings are robust when comparing them with the marginal effects from the probit estimations.²⁶

Table 10 reports the results on having income from work in LTPS. Column (1) and (2) present the results from OLS regressions with only treatment variables, and with treatment variables and covariates. Column (3) and (4) report the marginal effects from probit estimations. As the girls have graduated from Form IV by the time of the LTPS, we expect to find that more girls report having an income in LTPS. If this is true, we also might detect more clear differences between the four schools in LTPS than were found in STS.

²⁵ See Table B2: Income (STS) in Appendix B.

²⁶ Ibid.

Table 10: Income (LTPS)

	OLS (1)	OLS (2)	Probit (3)	Probit (4)
Health (d)	0.075 (0.086)	0.021 (0.092)	0.094 (0.100)	0.040 (0.105)
Business (d)	0.261*** (0.096)	0.230** (0.099)	0.291** (0.121)	0.266** (0.127)
Health & business (d)	0.094 (0.091)	0.129 (0.100)	0.117 (0.109)	0.171 (0.128)
Wealthy household (d)		0.017 (0.081)		0.021 (0.081)
Cognitive abilities (d)		-0.105 (0.068)		-0.112 (0.069)
Business knowledge		-0.011 (0.048)		-0.014 (0.048)
Age > 17 (d)		-0.071 (0.068)		-0.080 (0.067)
Risk averse (d)		-0.037 (0.069)		-0.033 (0.069)
Constant	0.128* (0.067)	0.248** (0.118)		
R^2	0.043	0.073		
t_HB	0.336	0.251	1.158	0.969
se_HB	0.160	0.167	0.583	0.657
Observations	179	151	179	151

Note: The table reports average treatment effects on having income from work, where the dependent variable is an indicator variable taking the value 1 if the girl had business plans in LTPS. Columns (1) and (2) report outcomes from OLS regressions and columns (3) and (4) reports marginal effects from probit estimations. "Health", "Business" and "Health & business" are treatment variables taking the value 1 if the girl was assigned to this treatment. t_HB and se_HB shows the point estimates and their belonging standard errors for a linear combination of the "Health" and "Business" coefficients. Differences in sample size when covariates are included reflect missing information from some of the participants. Marginal effects; Standard errors in parentheses, (d) for discrete change of dummy variable from 0 to 1, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

As expected from the findings in STS, there are more girls with income in LTPS. The only significantly positive long-term treatment effect found from the regressions is from the business training. The coefficient for the business training in column (1) indicates that a girl is 26 percentage points more likely to have income from work in the LTPS than a girl from the control group. Compared to a girl from the control group, the probability of having income from work almost triples if a girl has been offered business training. However, we do not observe a significant effect from the combined treatment. As we only have data from four schools in LTPS, this could be explained by specific differences in the school's characteristics. Another possible explanation to the dominant effect of the business treatment is that the combined treatment has caused a *crowding out* effect: receiving both treatments has had a lower effect on having an income than what would have been the case if the girl only received the business training. The crowding out effect was already envisioned by the Girl Empowerment Project Team at the planning stage of the interventions. If these findings also are observed in the expected long-term survey, this could imply that investing in only business training could be a more efficient use of funds if the goal is to improve the entrepreneurial mind-set among young women in Tanzania.

The treatment effect from the health training is still not significant, and neither is the effect from the combined program. As in STS, we do not find any significant correlation between income and the various background characteristics in LTPS. The coefficient for the health training becomes negligible, indicating that a girl is almost just as likely to have income if she received health training as if she did not receive any training at all.

The findings are robust when comparing them with the marginal effects from the probit estimations presented in column (3) and (4).

We have shown that it was mainly the business treated girls who reported having an income in LTPS. In an open-ended question in the interviews from LTPS, the girls were asked about what had happened to their plan to start a business from STS, if they had one. Across all treatments, some of the girls state that they do not have plans to start a business anymore, as they do not have the necessary capital or skills. Other girls facing similar constraints have simply postponed their business start-up while acquiring the money or skills needed to start a business. From the data it is evident that the business treated girls are more proactive in their approach. The business treated girls seem to be actively searching for ways to fulfill their business plans to a greater extent than the other girls in the sample. "*The business plan has*

not changed because I still have interest in it, that is why I am learning how to sow so that I will be able to start the business afterwards” states a business treated girl with plans to start a business as a tailor. Another business treated girl states that *“I want to sell first grade ladies’ clothes after getting capital from the business I am doing now”*. She is currently running a business together with her sister that gives her a weekly income of 70,000 TZS. The two examples identify business treated girls who are finding a way to solve the constraints they are facing. It is important to note that there are also girls from the business treated school who claim to no longer have business plans due to lack of capital or skills, and are waiting for the constraints to evaporate by e.g. receiving money from parents. However, compared to the other treatment groups, the business treated girls seem to be more actively searching for ways to acquire the necessary skills and capital, rather than passively waiting for problems to be solved. Based on the girls’ personal essays, Maalim (2014) finds that the business training has provided the girls with the necessary motivation to take control over business outcomes and increase the girls’ self-awareness. Her findings suggest that girls with an internal locus of control are more likely to be proactive, which makes them more likely to succeed in business. This is in line with our findings of business treated girls being more proactive, and to a larger extent overcoming constraints to starting a business.

The solution-oriented aspect observed among the business treated girls is also reflected in the girls’ moving pattern.²⁷ A few of the girls claim to have moved to stay with their mother or grandmother; others have moved to stay with their sister or aunt. As mentioned previously in the analysis, girls who move to their grandparents mostly do so to help out with household activities, while girls who go to stay with sisters or aunts do so in search for opportunities within business or education. According to Jacqueline Mgumia, youth often follow resources and family members, like moving to an aunt who runs a small business and might need some help or could help you start your own business. In Table 11 we present the business treated girls’ reasons for moving. Among the business treated girls who have moved, about half of them have moved to look for more opportunities, to do business, or helping out others with their business. As an example, one of the girls state: *“I have moved to stay with my sister so that we can do business together”*. Another girl claims her reason for moving to be *“... to look for opportunities - especially in business”*. To conclude, the business treated girls do not only move more than the other treatment groups – they also tend to move after opportunities.

²⁷ The treatment effect on girls’ moving pattern is shown in Table C2: Moving (LTPS) in Appendix C.

Table 11: Business treated girls' reasons for moving

Answer to question Q 6.2: "Why did you move?"	Answer to question Q 13.1 "Are you planning to start up a (new) business?"
<i>I came to stay with my aunt</i>	Yes
<i>To help my sister in her business while waiting to get money to go back to school</i>	Yes
<i>To do business</i>	No
<i>To see how can I continue with my life</i>	Yes
<i>She has moved to stay with her so that they can do business together</i>	Yes
<i>To stay with her sister</i>	Yes
<i>She has moved to stay with her sister</i>	No
<i>She has moved to come and stay with her mother</i>	No
<i>I have come to help my brother</i>	No
<i>To look for more opportunities</i>	No
<i>She has moved to stay with her aunt</i>	No
<i>To look for more opportunities – especially in business</i>	Yes
<i>She has moved to stay</i>	No
<i>She has moved to her grandmother for the purpose of looking something to do</i>	Yes
<i>To stay with my grandmother</i>	No
<i>To look for more opportunities</i>	Yes

Note: The table shows the answers to question 6.2 and 13.1 for all business treated girls that have moved since LTPS.

Jacqueline Mgumia emphasizes that different families have different attitudes towards business. One possible aspect is that wealthier households are more likely to have a better understanding of business, and thus be more supportive of their daughters engaging in self-employment. Another aspect is household obligations: as young women often depend on the household to support them, they may have to do what is expected of them in return. Hence, if there are young children or sick family members in the household, it is often the girls'

responsibility to look after them. According to Mgumia, this leads to time constraints. All in all, a supportive family is important for young entrepreneurs to succeed – especially for young women. It is possible that the families of business treated girls have become more supportive of business activities due to the training, and thus encourage or accept that their girls move toward opportunities. However, the treatment effect on moving could also be caused by the girls themselves actively searching for opportunities, which may include moving. As there is only one school from each treatment in the four school sample, there can be certain school characteristics that cause the rather large observed treatment effects.

Comparing the locations of the four schools in the four school sample,²⁸ it is evident that although the business treated school is categorized as a remote school, its location is not far from Morogoro, a city with a population of 300,000. In Morogoro, the roads to bigger cities like Dodoma, Dar es Salaam and Iringa meet, making the city a fairly large trading center. The other three schools in the four school sample are located close to each other – not far from the small rural town Ifakara. Due to the small sample in LTPS, with only one school from each treatment group, it is possible that the effect from business training on having an income is triggered by the business treated school's proximity to a large trading center. Youth in this area might have a different mindset and perception of business, compared to those from more remote schools. There could also be other characteristics of the business treated school that differentiates it from other schools in the four school sample, such as better nutrition through regular school meals, or particularly engaged and talented teachers that improve learning outcomes. The planned long-term survey will provide more robust findings to our research question, as it will contain data from the full randomized sample.

²⁸ See Figure 1

8 Conclusion

The main objective of this paper has been to provide and analyze updated information of a subsample of the girls involved in the Girl Empowerment Project, in order to say something about the girls' plans to start a business, their income-generating activities, and the challenges they have faced since graduating from secondary school. By comparing data from a long-term pilot survey to data collected in previous surveys, we provide implications for how the treatments have affected the girls' entrepreneurial behavior in the longer term.

We find that the business training has successfully encouraged the girls to develop business plans, and findings from the new data suggest that business training also has inspired them to commence income-generating activities.

The combined treatment seems to have had an even larger effect on plans to start a business than the business treatment alone. However, we do not find that the combined treatment increased the probability of girls having an income in the long-term pilot survey. We consider the possibility that receiving both the health and business training may have caused a crowding out effect.

Qualitative data from the long-term pilot survey show the tendency of business treated girls handling constraints in a more proactive fashion than girls from other treatment groups. By making an effort to acquire the money and skills needed to start a business, and moving towards business opportunities, these girls clearly stand out by having a more entrepreneurial mind-set.

However, as the sample examined in the long-term pilot survey does not fulfill the guidelines for achieving a representative sample, these findings cannot be generalized to the full sample. We also take note of the fact that the business school is located close to a large trading center, which could explain the business treated girls' entrepreneurial behavior.

This research adds to the literature by increasing the understanding of the effectiveness of business training on young women with no business experience prior to the training. Through the implementation of the long-term pilot survey, we found that both treated and untreated girls were easy to reach and willing to share information. These findings indicate that phone interviews can be a useful tool for collecting data in a developing country.

References

- Ackroyd, S., & Hughes, J. (1981). Data collection in context. *Longman*.
- Aker, J., & Mbiti, I. M. (2010). Mobile phones and economic development in Africa. *Journal of Economic Perspectives*, 207-232.
- Angrist, J. D., & Pischke, J.-S. (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, New Jersey: Princeton University Press.
- Angrist, J., & Pischke, J.-S. (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, New Jersey: Princeton University Press.
- Bailey, M. J. (2006). More power to the pill: the impact of contraceptive freedom on women's life cycle labor supply. *Quarterly Journal of Economics*, 121: 289-320.
- Banerjee, A., Duflo, E., Glennester, R., & Kinnan, C. (2013). *The miracle of microfinance? Evidence from a randomized evaluation*. NBER Working Paper 18950.
- Berge, L. I., Bjorvatn, K., & Tungodden, B. (2014). Human and financial capital for microenterprise development: Evidence from a field and lab experiment.
- Berge, L. I., Bjorvatn, K., Somville, V., & Tungodden, B. (2012). *Pre-Analysis Plan*.
- Binachi, A. (1993). Who's most likely to do it alone? *Inc.*, 15(5): 58.
- Bjorvatn, K., Cappelen, A., Sekei, L., Sørensen, E., & Tungodden, B. (2014). *Teaching through television: Experimental evidence on entrepreneurship education in Tanzania*. Bergen, Norway: Working paper. NHH Norwegian School of Economics.
- Bruhn, M., & Zia, B. (2011). *Stimulating managerial capital in emerging markets: The impact of business and financial literacy for young entrepreneurs*, *The World Bank Working Paper*. The World Bank, Working Paper.
- Business Dictionary. (n.d.). *Business Dictionary*. Retrieved 06 2014, from <http://www.businessdictionary.com/definition/formal-sector.html>
- Byers, T., Kist, H., & Sutton, R. (1997). *Characteristics of the entrepreneur: Social creatures, not solo heroes*. Boca Raton, FL.: CRC Press LLC.

- CIA. (2013). *World Fact Book*. Retrieved 03 25, 2014, from Central Intelligence Agency: <https://www.cia.gov/library/publications/the-world-factbook/geos/tz.html>
- Collins, D., Murdoch, J., Rutherford, S., & Ruthven, Q. (2009). *Portfolios of the poor: How the world's poor live on \$2 a day*. Princeton, NJ: Princeton University Press.
- Daily News. (2014, May 23). *Thousand fail to pay exam fees*. Retrieved 06 12, 2014, from <http://m.dailynews.co.tz/index.php/local-news/31728-thousands-fail-to-pay-exam-fees>
- de Mel, S., McKenzie, D., & Woodruff, C. (2008). Returns to capital in microenterprises: Evidence from a field experiment. *Quarterly Journal of Economics*, 123 (4): 1329-1371.
- de Mel, S., McKenzie, D., & Woodruff, C. (2012b). Business training and female enterprise start-up, growth, and dynamics: Experimental evidence from Sri Lanka.
- Deaton, A. (2010). Instruments, randomization, and learning about development. *Journal of Economic Literature*, 425-455.
- Djankov, S., Lieberman, I., Mukherjee, J., & Nenova, T. (2002). *Going Informal: Benefits and Costs*. Draft World Bank.
- Drageset, S., & Ellingsen, S. (2009). Forståelsen av kvantitativ helseforsikring - en introduksjon og oversikt. *Nordisk Tidsskrift for Helseforskning*, 100-113.
- Duflo, E., Glennerster, R., & Kremer, M. (2006). Using randomization in development economic research: A toolkit. *Massachusetts Institute of Technology*.
- Dupas, P. (2011a). Health behavior in developing countries. *Annual Review of Economics*, 3: 425-449.
- Dupas, P. (2011b). Do teenagers respond to HIV risk information? Evidence from a field experiment in Kenya. *American Economic Journal: Applied Economics*, 3(1), pp. 1–34.
- ESRF. (2012). *Informal sector taxation in Tanzania*. Retrieved 05 15, 2014, from Economic and Social Research Foundation: http://esrf.or.tz/docs/INFORMATIONSECTORTAXATIONINTANZANIA_12.PDF

-
- Gine, X., & Mansuri, G. (2011). *Money or ideas? A field experiment on constraints to entrepreneurship in rural Pakistan*. The World Bank.
- Goldin, C., & Katz, L. (2000). The power of the pill: Oral contraceptives and women's career and marriage decision. *National Bureau of Economic Research*.
- Government of Tanzania. (2011). *The Tanzania Five Year Development Plan for 2011/2012-2015/2016 - Unleashing Tanzania's Latent Growth Potentials*. Dar es Salaam.
- Helgesson, L. (2006). *Getting ready for life*. Umeå: Umeå: Print & Media.
- Hobson, E. W. (2011). *The importance of the informal economy for local economic development (LED) in Africa*. Retrieved 03 27, 2014, from LEDNA: http://ledna.org/sites/ledna.org/files/the_importance_of_the_informal_economy_for_local_economic_development_in_africa.pdf
- IMF. (n.d.). *Chapter 7 - Rural microfinance institutions: Regulatory and supervisory issues*. Retrieved 05 30, 2014, from International Monetary Fund: <http://www.imf.org/external/pubs/ft/fsa/eng/pdf/ch07.pdf>
- IPP Media. (2014, 02 22). *Form four exams: Girls outdo boys again in top honours list*. Retrieved 05 25, 2014, from <http://www.ippmedia.com/frontend/?l=65074>
- Ishengoma, E., & Kappel, R. (2006). *Economic growth and poverty: Does formalisation of informal enterprises matter?* Hamburg, GER.
- Karlan, D., & Morduch, J. (2009). Access to finance, in Danr Rodrick and Marck Rosenzweig, eds. *Handbook of Development Economics*.
- Klinger, B., & Schündeln, M. (2007). *Can entrepreneurial activity be taught? Quasi-experimental evidence from Central America*. Hentet fra http://www.hks.harvard.edu/var/ezp_site/storage/fckeditor/file/pdfs/centers-programs/centers/cid/publications/faculty/wp/153.pdf
- Leonard, W., & Mfaume, R. M. (2004). *Small business entrepreneurship in Dar es Salaam - Tanzania: Exploring problems and prospects for future development*. Somerset West, RSA: Forum Paper.

- Maalim, A. M. (2014). Master thesis in International Business. Bergen, Norway: Norwegian School of Economics.
- Miller, G. (2010). Contraception as development? New evidence from family planning in Colombia. *Economic Journal*, 120:709-736.
- Ministry of Education and Vocational Training. (2012). *Basic education statistics in Tanzania 2012*. Retrieved 05 15, 2014, from http://www.moe.go/tz/index.php?searchword=BEST&ordering=&searchphrase=all&Itemid=564&option=com_search
- Muthengi-Karei, E., & Erulkar, A. (2012). *Increasing opportunities to delay marriage and promote schooling*. Dar es Salaam, Tanzania.
- National Bureau of Statistics. (2011). *Tanzania 2010 Demographic and Health Survey Key Findings*. Calverton, Maryland, USA.
- NECTA. (n.d.). *Certificate of secondary education examination (CSEE)*. Retrieved 05 15, 2014, from National Examinations Council of Tanzania: http://www.matokeo.necta.go.tz/necta2/exam_csee.html
- NECTA. (n.d.). *Registration matters*. Retrieved 05 12, 2014, from National Examination Council of Tanzania: <http://www.necta.go.tz/registrations.html>
- Opdenakker, R. (2006). Advantages and disadvantages of four interview techniques in qualitative research. *Qualitative Social Research*.
- Pearsall, J. (2002). Cognition. I *Concise Oxford English Dictionary*. Oxford, England: Oxford University Press.
- Popper, K. (1959). *The logic of scientific discovery*. Routledge, Taylor & Francis.
- Poverty Action Lab. (n.d.). *What is randomization?* Retrieved May 2014, from <http://www.povertyactionlab.org/methodology/what-randomization>
- Purdon, S., & Thomes, R. (1994). *Telephone methods for social surveys*. Surrey: University of Surrey.
- Rasul, I. (2008). Household bargaining over fertility: Theory and evidence from Malaysia. *Journal of Development Economics*, 86: 215-241.

-
- Schafer, D., & Ramsey, F. (2002). *The Statistical Sleuth: A course in methods of data analysis*. Duxbury Press.
- Stock, J. H., & Watson, M. M. (2012). *Introduction to Econometrics, Third Edition*. Essex, England: Pearson.
- Sweet, L. (2002). Telephone interviewing: Is it compatible with interpretive phenomenological research? *Contemporary Nurse*, 58-63.
- TCCIA. (2009). *Understanding entrepreneurship*. Retrieved 03 27, 2014, from Tanzania Chamber of Commerce, Industry and Agriculture:
<http://www.tccia.com/tcciaweb/SMEtoolkit/introduction.htm>
- Timmons, J. A. (1994). *New Venture Creation: Entrepreneurship for the 21st century*. . Burr Ridge, IL.: Irwin Press.
- TradeMark Southern Africa. (2013, 01 03). *Infrastructure and Tanzania's Long-Term Perspective Plan 2011/12-2025/26*. Hentet 06 14, 2014 fra
<http://www.trademarksa.org/news/infrastructure-and-tanzanias-long-term-perspective-plan-201112-202526>
- Tungodden, B. (2012). *Girl's economic empowerment – the best contraceptive? A randomized controlled trial in Tanzania. Application document*. Bergen, Norway.
- UN. (2013). *We Can End Poverty - Millennium Development Goals and Beyond 2015*. Retrieved from UN: <http://www.un.org/millenniumgoals/education.shtml>
- UN. (2013, September). *We Can End Poverty - Millennium Development Goals and Beyond 2015 Fact Sheet*. Retrieved from UN:
http://www.un.org/millenniumgoals/pdf/Goal_2_fs.pdf
- UN Tanzania. (n.d.). *Education*. Retrieved 03 27, 2014, from United Nations Tanzania:
<http://tz.one.un.org/index.php./what-we-do/education>
- UNFPA. (2013). *State of the world population 2013: Motherhood into childhood*. Retrieved 06 02, 2014, from United Nations Populations Fund:
<http://www.unfpa.org/webdav/site/global/shared/swp2013/EN-SWOP2013-final.pdf>

UNICEF. (n.d.). *Education in Tanzania*. Retrieved 03 26, 2014, from UNICEF Tanzania:
<http://www.unicef.org/tanzania/6911.html>

UNICEF. (n.d.). *Education in Tanzania*. Retrieved 03 26, 2014, from UNICEF Tanzania:
<http://www.unicef.org/tanzania/6911.html>

US Aid. (2013). *Tanzania education strategy for improving the quality of education 2009-2013*. Dar es Salaam.

WHO. (2012, 10 30). *Early marriages, adolescent and young pregnancies*. Retrieved 06 02, 2014, from World Health Organization:
http://apps.who.int/gb/ebwha/pdf_files/WHA65/A65_13-en.pdf?ua=1

WHO. (n.d.). *Adolescent pregnancy*. Retrieved 05 30, 2014, from World Health Organization:
<http://www.who.int/mediacentre/factsheet/fs364/en/>

World Bank. (2013). *Tanzania Overview*. Retrieved 03 26, 2014, from World Bank:
<http://www.worldbank.org/en/country/tanzania/overview#1>

List of Figures

Figure 1: Distribution of schools.....	24
Figure 2: Calls per girl	37
Figure 3: Constraints to starting a business.....	47

List of Tables

Table 1: Timeline of the project.....	26
Table 2: Call history.....	36
Table 3: Family status and education.....	41
Table 4: Regular income and farming.....	43
Table 5: Business plans STS and LTPS.....	45
Table 6: Business plans STS and income LTPS.....	45
Table 7: Business plans (STS).....	51
Table 8: Business plans (LTPS).....	53
Table 9: Income (STS).....	57
Table 10: Income (LTPS).....	59
Table 11: Business treated girls' reasons for moving.....	62

Abbreviations

AIDS	Acquired Immunodeficiency Syndrome
BEST	Basic Education Statistics in Tanzania
CIA	Central Intelligence Agency
CMI	Christian Michelsen Institute
CSEE	Certificate of Secondary Education Examinations
DPC	Development Pioneer Consultants LTD
ESRF	Economic and Social Research Foundation
GDP	Gross Domestic Product
HIV	Human Immunodeficiency Virus
IMF	International Monetary Fund
LTPS	Long-term pilot survey
NECTA	National Examinations Council of Tanzania
NHH	Norwegian School of Economics
OLS	Ordinary Least Squares
PRIDE	Promotion of Rural Initiatives and Development Enterprise
PSLE	Primary School Leaving Examination
UDEC	University of Dar es Salaam Entrepreneurship Centre
UN	United Nations
UNFPA	United Nations Population Fund
UNICEF	United Nations' Children's Fund
US Aid	United States Agency for International Development
RCT	Randomized Controlled Trial
STS	Short-term survey
TCCIA	Tanzania Chamber of Commerce, Industry and Agriculture
TZS	Tanzanian shillings
WHO	World Health Organization

Appendix A: Interview Guide for Phone Interviews LTP

Dear _____ my name is _____ and I was part of the team which conducted a study about young women in Tanzania which you participated in when you were in Form IV in 2013. As you may recall, researchers came to your school and asked you to fill in a questionnaire. After six months we came back to your school and we conducted another round of questionnaires where you answered questions and could earn money. Do you remember this? At this point we would like to know how the girls in this project are doing, and we therefore would like to ask you a few questions over the phone. Is this OK with you? Your responses to the questions in the survey are private, and will only be seen by the researchers. You may find some questions difficult, but please do your best and take all the time you need before you answer. If you have any questions now or during this session, feel free to ask.

I CONTACT INFORMATION

Background information from baseline (fill in the ID before calling):

Q1	ID:	
Q2	Date of interview:	[EXTRACT FROM TIME VARIABLES]
Q3	Surveyor:	

Q4_1	Name (full name)	Q4_2	If changed:
	[FROM DATA]		
Q5_1	Name of household head (father, mother or guardian):	Q5_2	If changed:
	[FROM DATA]		

Q6_1	Have you moved since the last time you participated in our survey (April/October)?	(Y/N)
------	--	-------

	<i>If YES:</i>	
Q6_2	Why did you move?	
Q6_3	Name of new village:	
Q6_4	New landmark:	
Q6_5	Ward where you live now:	

II BACKGROUND INFORMATION

Q7	Have you passed the exam for Form IV?	(Y/N)
----	---------------------------------------	-------

Q8	What is your main daily activity these days? Describe.	
----	--	--

Q9_1	Are you married?	(Y/N)
------	------------------	-------

Q9_2	How many children do you have?	
------	--------------------------------	--

Q9_3	Are you pregnant?	(Y/N)
------	-------------------	-------

III INCOME

Q10_1	Do you have a regular income?	(Y/N)
-------	-------------------------------	-------

Q11_1	If NO, what are you doing?	
	Misha-misha (looking for a job)	
	Nipo-nipo (nothing)	
Q11_2	Other (please specify):	

	If YES, what is the source of income (multiple sources possible)?	Please specify what kind of source?		What is the income from this source in a normal week? (TZS)
Q12_1_1	Employment		Q12_1_2	
Q12_2_1	Own business		Q12_2_2	
Q12_3_1	Other		Q12_3_2	

IV BUSINESS PLANS (also ask those who already have a business)

Q13_1	Are you planning to start up a (new) business?	(Y/N)
Q13_2	If yes, please specify:	

Q14_1	In the previous survey, you wrote that your plan was...	[FROM DATA]
Q14_2	How do you see this plan today? Has it changed? Why/why not? Describe.	

V CONSTRAINTS TO STARTING UP OR RUNNING A BUSINESS

Rank the following constraints to starting up or running a business, as you see it, where 1 means the most important constraint and 4 means the least important constraint:

	Constraint	Rank		Specify
Q15_1_1	Family obligations		Q15_1_2	
Q15_2_1	Start-up capital		Q15_2_2	
Q15_3_1	Business idea		Q15_3_2	
Q15_4_1	Other		Q15_4_2	

VI FINAL REMARKS

Q16	Is there anything you would like to add?	
-----	--	--

Thank you for your participation. If you have questions or additional information, do not hesitate to call me at any time on my mobile phone number.

VII EVALUATION OF PHONE CALL BY RESEARCH ASSISTANT

Q17	How easy was it to reach this girl?				
	Very difficult	Somewhat difficult	Neutral	Somewhat easy	Very easy

Q18	How willing was the girl to share the information you asked for?				
	Very unwilling	Somewhat unwilling	Neutral	Somewhat willing	Very willing

Appendix B: Probit estimations

Table B1: Plans to start a business (STS)

	Full sample		Remote sample		Four school sample	
	(1)	(2)	(3)	(4)	(5)	(6)
Health (d)	0.013 (0.045)	0.027 (0.043)	0.109** (0.051)	0.106** (0.045)	0.327* (0.168)	0.356* (0.192)
Business (d)	0.393*** (0.044)	0.411*** (0.042)	0.445*** (0.037)	0.452*** (0.036)	0.662*** (0.121)	0.764*** (0.111)
Health & business (d)	0.428*** (0.047)	0.461*** (0.046)	0.548*** (0.053)	0.568*** (0.046)	0.800*** (0.082)	0.832*** (0.086)
Wealthy household (d)		0.029 (0.025)		0.044 (0.032)		0.173 (0.119)
Cognitive abilities (d)		0.060*** (0.020)		0.040* (0.023)		0.134 (0.082)
Business knowledge		0.045*** (0.013)		0.039** (0.016)		-0.020 (0.056)
Age > 17 (d)		0.006 (0.019)		0.001 (0.027)		-0.138 (0.085)
Risk averse (d)		0.017 (0.019)		0.019 (0.025)		0.035 (0.087)
t_HB	1.077	1.170	1.520	1.537	2.822	3.430
se_HB	0.213	0.202	0.220	0.203	0.908	1.058
Observations	3305	2849	1691	1459	166	138

Note: The table reports marginal effects from probit estimations on the average treatment effects on plans to start a business, where the dependent variable is an indicator variable taking the value 1 if the girl had business plans in the short-term survey. Columns (1) and (2) report results for the full sample. Columns (3) and (4) report outcomes for only remote schools. Columns (5) and (6) report results for the four school sample. "Health", "Business" and "Health & business" are treatment variables taking the value 1 if the girl was assigned to this treatment. t_HB and se_HB shows the point estimates and their belonging standard errors for a linear combination of the "Health" and "Business" coefficients. Differences in sample size when covariates are included reflect missing information from some of the participants. Marginal effects; Standard errors in parentheses, clustered at school level for the remote and full sample, (d) for discrete change of dummy variable from 0 to 1, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B2: Income (STS)

	Full sample		Remote sample		Four school sample	
	(1)	(2)	(3)	(4)	(5)	(6)
Health (d)	-0.005 (0.012)	-0.006 (0.011)	0.012 (0.017)	0.023 (0.016)	0.494 (54.356)	0.568 (106.671)
Business (d)	0.035** (0.015)	0.033** (0.015)	0.035** (0.015)	0.042*** (0.015)	0.799 (45.655)	0.781 (86.405)
Health & business (d)	0.035** (0.017)	0.040* (0.020)	0.041* (0.021)	0.050** (0.022)	0.799 (42.714)	0.721 (101.127)
Wealthy household (d)		0.010 (0.010)		0.002 (0.011)		0.007 (2.834)
Cognitive abilities (d)		0.003 (0.009)		-0.007 (0.012)		0.001 (0.653)
Business knowledge		0.003 (0.005)		-0.000 (0.006)		-0.000 (0.090)
Age > 17 (d)		0.008 (0.009)		0.004 (0.010)		-0.001 (0.386)
Risk averse (d)		-0.000 (0.008)		-0.002 (0.012)		-0.001 (0.536)
t_HB	0.241	0.220	0.463	0.641	7.767	9.035
se_HB	0.183	0.190	0.226	0.229	575.350	1169.554
Observations	3304	2849	1690	1459	165	138

Note: The table reports marginal effects from probit estimations on the average treatment effects on having income from work where the dependent variable is an indicator variable taking the value 1 if the girl had business plans in the short-term survey. Columns (1) and (2) report results for the full sample. Columns (3) and (4) report outcomes for only remote schools. Columns (5) and (6) report results for the four school sample. "Health", "Business" and "Health & business" are treatment variables taking the value 1 if the girl was assigned to this treatment. t_HB and se_HB shows the point estimates and their belonging standard errors for a linear combination of the "Health" and "Business" coefficients. Differences in sample size when covariates are included reflect missing information from some of the participants. Marginal effects; Standard errors in parentheses, clustered at school level for the remote and full sample, (d) for discrete change of dummy variable from 0 to 1, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix C: Marriage, moving and treatment effects

Table C1: Marriage and permanent relationships (LTPS)

	OLS (1)	OLS (2)	Probit (3)	Probit (4)
Health (d)	0.173* (0.095)	0.045 (0.107)	0.231* (0.125)	0.060 (0.119)
Business (d)	0.090 (0.175)	0.072 (0.179)	0.145 (0.240)	0.067 (0.222)
Health & business (d)	0.212** (0.087)	0.146 (0.094)	0.254** (0.106)	0.156 (0.111)
Wealthy household (d)		0.021 (0.087)		0.014 (0.083)
Cognitive abilities (d)		-0.073 (0.078)		-0.089 (0.077)
Business knowledge		-0.067 (0.058)		-0.076 (0.054)
Age > 17 (d)		-0.017 (0.077)		-0.011 (0.071)
Risk averse (d)		-0.206** (0.082)		-0.209** (0.085)
Constant	0.077 (0.064)	0.351*** (0.132)		
R^2	0.052	0.132		
t_HB	0.263	0.117	1.210	0.536
se_HB	0.219	0.228	0.882	1.061
Observations	122	97	122	97

Note: The table reports average treatment effects on being married or in a permanent relationship, where the dependent variable is an indicator variable taking the value 1 if the girl had business plans in the LTPS. Columns (1) and (2) report outcomes from OLS regressions and columns (3) and (4) reports marginal effects from probit estimations. "Health", "Business" and "Health & business" are treatment variables taking the value 1 if the girl was assigned to this treatment. t_HB and se_HB shows the point estimates and their belonging standard errors for a linear combination of the "Health" and "Business" coefficients. Differences in sample size when covariates are included reflect missing information from some of the participants. Marginal effects; Standard errors in parentheses, (d) for discrete change of dummy variable from 0 to 1, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table C2: Moving (LTPS)

	OLS (1)	OLS (2)	Probit (3)	Probit (4)
Health (d)	-0.011 (0.095)	0.001 (0.102)	-0.011 (0.097)	-0.003 (0.108)
Business (d)	0.190* (0.107)	0.218** (0.110)	0.189* (0.115)	0.235* (0.123)
Health & business (d)	-0.015 (0.101)	0.120 (0.112)	-0.016 (0.102)	0.128 (0.126)
Wealthy household (d)		-0.122 (0.090)		-0.118 (0.085)
Cognitive abilities (d)		-0.061 (0.075)		-0.070 (0.079)
Business knowledge		-0.112** (0.053)		-0.119** (0.056)
Age > 17 (d)		-0.102 (0.076)		-0.103 (0.077)
Risk averse (d)		0.138* (0.077)		0.146* (0.079)
Constant	0.282*** (0.074)	0.418*** (0.132)		
R^2	0.030	0.104		
t_HB	0.179	0.219	0.475	0.630
se_HB	0.177	0.186	0.506	0.561
Observations	179	151	179	151

Note: The table reports average treatment effects on moving, where the dependent variable is an indicator variable taking the value 1 if the girl had business plans in the LTPS. Columns (1) and (2) report outcomes from OLS regressions and columns (3) and (4) reports marginal effects from probit estimations. "Health", "Business" and "Health & business" are treatment variables taking the value 1 if the girl was assigned to this treatment. t_HB and se_HB shows the point estimates and their belonging standard errors for a linear combination of the "Health" and "Business" coefficients. Differences in sample size when covariates are included reflect missing information from some of the participants. Marginal effects; Standard errors in parentheses, (d) for discrete change of dummy variable from 0 to 1, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix D: Business knowledge questions from baseline survey

3.1.1 Which of the following is an important part of customer service? (Tick off one.)

- A To never recommend the most expensive products to customers
- B To always praise the goods you sell
- C To be reliable in relations with the customer
- D To always recommend cheap products to customers

3.1.2 What is profit? (Tick off one.)

- A Profit is sales of the most important products
- B Profit is sales minus cost of goods and operating expenses
- C Profit is sales minus cost of goods and what you take home from the business
- D Profit is sales plus cost of goods and operating expenses

3.1.3 Why is it important for the business to keep stock? (Tick off one.)

- A To have goods available for family consumption
- B To have goods available for the customers
- C To have goods available for the suppliers
- D To have goods available for a family emergency