Charles University in Prague

Faculty of Social Sciences
Institute of Economic Studies

BACHELOR THESIS

Ghana cocoa farmers' survey: chocolate economics

Author: **Michal Svatoň**

Supervisor: **Petr Janský M.Sc.**

Academic Year: **2011/2012**

Declaration of Authorship

I hereby declare that I have written this thesis using only literature and other sources listed in reference list. Furthermore, I declare that I have not used this thesis to acquire another academic degree. I acknowledge and agree with lending and publishing of the thesis.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis document in whole or in part.

Prague, May 15, 2012

Signature

Acknowledgments

The author is grateful to Petr Janský M.Sc. for his useful comments and patient supervision and to prof. RNDr. Jan Ámos Víšek CSc. for his advice. My special thanks go to Center for the Study of African Economies, Oxford for providing me with the data.

Abstract

In this thesis I evaluate long-term productive and economic impacts of the Cocoa Abrabopa Association, a private sector initiative providing fertilizer and other agrochemicals to Ghanaian cocoa farmers on credit with a joint liability. Both the productive and economic returns of the program are expected to be higher in the latter years of sustained membership in the program due to the accumulation of nutrients in the soil and the possibility of the extension of the size of the loan provided by the program. Earlier work focused only on the estimation of the effect of adoption of agricultural technology but did not allow to examine possible difference in returns of a longer participation. This thesis contributes by evaluating the impacts in the 1st, the 2nd and the 3rd year of the CAA membership and their comparison. By constructing and estimating a model for the impacts of longer participation I confirm the hypothesis of higher economic returns during the 2nd and the 3rd year of participation in the CAA program. The evidence which would confirm the hypothesis of higher agronomic returns has not been found in the case of CAA.

JEL Classification	C21, C23, Q16, Q18
Keywords	Cocoa Abrabopa Association, Ghana, economic return, long run, microcredit, fertilizer

Author's e-mail	`svaton.michal@gmail.com`
Supervisor's e-mail	`jansky@fsv.cuni.cz`

Abstrakt

V této práci hodnotím agronomické i ekonomické dlouhobodé dopady Cocoa Abrabopa Association - soukromé iniciativy poskytující hnojiva a další agrochemikálie ghanským kakaovým farmářům formou úvěru se společným ručením. Kvůli akumulaci živin v půdě během soustavného užívání hnojiv a také možnosti zvětšení úvěru lze očekávat, že zemědělská i ekonomická návratnost v pozdějších letech soustavné účasti předčí návratnost v prvním roce. Dřívější literatura se zaměřila na odhadování dopadu účasti v programu, avšak nezkoumala možné rozdíly v návratnosti v závislosti na délce účasti v programu. Tato práce přispívá studiem dopadů účasti po dobu 1, 2 a 3 roků a jejich vzájemným porovnáním. Zkonstruováním a odhadnutím modelu pro dlouhodobé efekty programu potvrzuji hypotézu vyšší ekonomické návratnosti v 2. a 3. roce účasti v programu. Hypotézu vyšší zemědělské návratnosti programu Cocoa Abrabopa Association tato práce nepodporuje.

Klasifikace JEL	C21, C23, Q16, Q18
Klíčová slova	Cocoa Abrabopa Association, Ghana, návratnost, dlouhé období, mikrokredit, hnojivo

E-mail autora	svaton.michal@gmail.com
E-mail vedoucího práce	jansky@fsv.cuni.cz

Contents

List of Tables viii

List of Figures ix

Acronyms x

Thesis Proposal xi

1 Introduction 1

2 Background of the study 3

2.1 Ghanaian cocoa sector 3
2.2 Program description 6
2.3 Sample description 7

3 Related work and problem statement 13

3.1 Literature review 13
3.2 Problem statement 15

4 Evaluation of the impact of the program on output 18

4.1 Output overview 18
4.2 Basic model specification 19
4.3 Extended model specification 21
4.4 Estimation and results 25

5 Evaluation of economic impacts 28

5.1 Cost-benefit mechanism 28
5.2 Non-labor inputs 29
5.3 Labor inputs 32
5.4 Economic returns computation 34

6 Retention **36**
6.1 Retention overview 36
6.2 Empirical evidence 37

7 Conclusion **40**

Bibliography **43**

A Appendix **I**
A.1 Tables and figures I
A.2 Cleaning procedure - data on received dose IV

List of Tables

4.1 Output average treatment effect estimates 25

5.1 Input usage - fraction of farmers 29
5.2 Average and median input quantities - non-members 30
5.3 Average and median input quantities - members 30
5.4 Demand estimates - non-labor inputs 31
5.5 Labor demand - in days 33
5.6 Labor demand 33
5.7 Demand estimates - labor inputs 34

6.1 Program retention estimates - probit model 38

A.1 Age distribution I
A.2 Total production and yield evolution II
A.3 Output estimation results III
A.4 Output ATE, only early adopters, one outlier removed III
A.5 Output ATE, yield $\in (30; 2000)$ III

List of Figures

2.1 Share of agriculture and cocoa production on Ghanaian GDP . . 4
2.2 Evolution of overall cocoa production 4
2.3 Credit share in cocoa sector 5
2.4 Age distribution in the sample 8
2.5 Female share in the sample 9
2.6 Farm size distribution 9
2.7 Education of nonmembers 10
2.8 Education of members 11
2.9 Hybrid trees usage 12

4.1 Output by seasons and treatment 18
4.2 Yield by seasons and treatment 19

A.1 Map of Ghanaian regions I
A.2 Regional share in cocoa production II

Acronyms

GDP Gross Domestic Product

CAA Cocoa Abrabopa Association

CRIG Cocoa Research Institute of Ghana

COCOBOD Ghana Cocoa Board

CSAE Center for the Study of African Economies, Oxford

ATE Average Treatment Effect

ROR Rate of return

JSS Junior Secondary School

Bachelor Thesis Proposal

Author	Michal Svatoň
Supervisor	Petr Janský M.Sc.
Proposed topic	Ghana cocoa farmers' survey: chocolate economics

Preliminary scope of work This work focuses on the evaluation of agronomic and economic impact of Cocoa Abrabopa Association programme which provides packages of fertilizer and other high-tech inputs on credit to cocoa farmers in Ghana. Contrary to high expected returns to treatment, the programme suffers from high drop-out rate of the members. As was explored by Zeitlin, Caria, Dzene, Janský, Opoku & Teal (2010), one of the key factors influencing drop-out decisions is the heterogeneity in returns. My goal is to explore the role of transient and persistent heterogeneity in returns for drop-out from Cocoa Abrabopa Association programme. The distinction between persistent and transient nature of heterogeneity in returns will have important consequences for policy of the programme in the sense of advocating widespread adoption and the sustainability of the programme itself.

Předběžná náplň práce Cílem této práce je vyhodnotit agronomické a ekonomické důsledky programu Cocoa Abrabopa Association, který poskytuje ghanským kakaovým farmářům na úvěr balíčky obsahující hnojiva a další high-tech vstupy. Tento program, ačkoliv vede k vysoké očekávané návratnosti, je doprovázen překvapivě vysokým množstvím účastníků opouštějících program. Jak bylo prozkoumáno v Zeitlin, Caria, Dzene, Janský, Opoku & Teal (2010), jedním z klíčových faktorů ovlivňujících rozhodování o opuštění programu je heterogenita výnosů. Mým cílem je prozkoumat roli dočasných a trvalých vlivů na heterogenitu výnosů pro rozhodování o opuštění programu. Odlišení trvalých a dočasných vlivů by mělo v případě významné role trvalé heterogenity důležité důsledky pro řízení programu – všeobecné šíření programu by se stalo těžce obhajitelným.

Outline

1. Introduction
2. Background of the study
3. Related work
4. Evaluation of the impacts
5. Conclusion

Core bibliography

1. CARIA, S., et al. (2009). "Impacts of group-based microfinance in agriculture : Evidence from Ghana." *Paper presented at the CSAE annual conference 2009.*
2. Zeitlin, A., Caria, S., and Teal, F. (2009). "Improving productivity through group lending: Report on the impact evaluation of the cocoa abrabopa initiative." *CSAE Working Paper.*
3. Zeitlin, A., Caria, S., Teal, F., Dzene, R., Jansky, P., and Opoku, E. (2010). "Heterogeneous returns and the persistence of agricultural technology adop- tion." *CSAE Working Paper.*
4. World Bank (2008), "World Development Report 2008: Agriculture for Development, The International Bank for Reconstruction and Development", Washington, D.C.
5. Zeitlin, A. (2009) "Producer networks and technology adoption in Ghana." *Paper presented at Midwest International Economic Development Conference.*

Author

Supervisor

Chapter 1

Introduction

In many developing countries agriculture represents a substantial share of GDP and more importantly provides employment to a huge part of the population. In contrast to its importance the productivity of agriculture is substantially lower than in developed countries. Low usage of hi-tech agricultural inputs, namely hybrid varieties, fertilizer and other agronomic chemicals, along with high labor intensiveness are considered to be some of the main causes of low productivity of agriculture in developing countries. As a result, it seems necessary to increase the adoption rates of agricultural technology in order to improve the living standards of farmers.

It is no surprise that Ghana fits the foreshadowed picture. Its share of agriculture on GDP is still large, though it has been declining. In particular, the production and marketing of cocoa beans represent 3% of Ghanaian GDP. In comparison with its main competitors, its productivity is low which is attributed mainly to the low usage of agricultural technology. Besides many other reasons for low adoption rates of these technologies there is one, which has received a lot of attention, and it is the credit constraint farmers are suffering from. Under this constraint many farmers are unable to invest into agricultural technology despite the high returns associated with it.

To alleviate the credit constraint for cocoa farmers Cocoa Abrabopa Association (CAA) program was launched. This program provides hi-tech agricultural inputs along with input application and business training on credit to groups of farmers with a joint liability. The program seems to be successful based on the rapidly increasing number of members of the program - from 1,440 farmers in 2006/2007 season to more than 19,000 farmers nowadays.

Zeitlin et al. (2009b), Zeitlin et al. (2009a) and Zeitlin et al. (2010) estimated high average positive economic returns on program participation which can be considered as another sign of success of the program. Nevertheless, this high positive return is accompanied by substantial heterogeneity in returns (Zeitlin et al. 2010) and surprisingly high drop-out rates. High drop-out rates could be partially attributed to the low individual returns of a portion of farmers. Still, low retention rate presents a serious problem. This problem will become even more serious if the returns during sustained membership exceed those in the 1st year as leaving the program will prevent farmers from a enjoying the benefits of staying in the program. Nevertheless, earlier literature did not investigate whether the difference is present.

The objective of this thesis is to identify the impacts of the CAA program not only during first (as in earlier literature), but also during 2nd a 3rd year of sustained membership and their comparison. To do so, this thesis introduces the extended identification strategy which will allow us to estimate the impacts of the program in 2nd and 3rd year of membership. Due to the possibility of the extension of the loan size in the 2nd and the 3rd of membership and higher productive impact of fertilizer usage in the long run (Hartemink 2005) both the productive and economic returns are expected to be higher for farmers staying in the program for more years. This thesis finds evidence for higher economic returns in latter years. Evidence for higher agronomic returns has not been found.

This thesis is structured in the following way: Chapter 2 describes Ghanaian cocoa sector, CAA program and the data in detail, chapter 3 presents an overview of related literature and discusses the objective of this thesis. Chapter 4 describes the model used for the empirical testing of the hypothesis and provides the estimates of productive impacts. Chapter 5 explains the cost benefit mechanism and provides the estimates of economic returns on CAA program participation. Chapter 6 covers the model of program retention. Chapter 7 concludes.

Chapter 2

Background of the study

2.1 Ghanaian cocoa sector

In Ghana, like in many other developing countries, agriculture constitutes a vital part of GDP. Even though its share of GDP has been declining during the last decade, it still represents about 30% of GDP and even more importantly about half of the employment. Cocoa production and marketing, the sub-sector of interest of this thesis, represents itself about 3% of GDP. The evolution of GDP shares over the examined period is displayed in figure 2.1.[1] Furthermore, cocoa along with gold is one of the biggest export articles of Ghana (Bank of Ghana, Ministry of Finance and Economic Planning).
Historically, Ghana was the largest producer of cocoa over major part of 20th century. Ghana lost its leading position in the season 1977/1978 to Côte d' Ivoire as mentioned in Quartey (2007). Still, western Africa is the leading region in cocoa production followed by southeast Asian countries (Indonesia,. . .) and South America (ICCO 2010). The evolution of Ghanaian cocoa sector output is displayed in figure 2.2.

Even though the Ghanaian cocoa sector production has been overall increasing, what is alarming is the low average yield of cocoa in Ghana over the past years (400 kg/ha) in comparison with its main competitors: Côte d' Ivoire (600 kg/ha) and Indonesia (800 kg/ha). This cannot be attributed to any country specific disadvantage as a yield close to 2500 kg/ha was reached on an experimental field by Cocoa Research Institute of Ghana (1973). Such a low yield

[1] The rationale of period selection is the availability of the data and interference of this period with waves of CAA program. This issue will be examined in the following section.

Figure 2.1: Share of agriculture and cocoa production on Ghanaian GDP

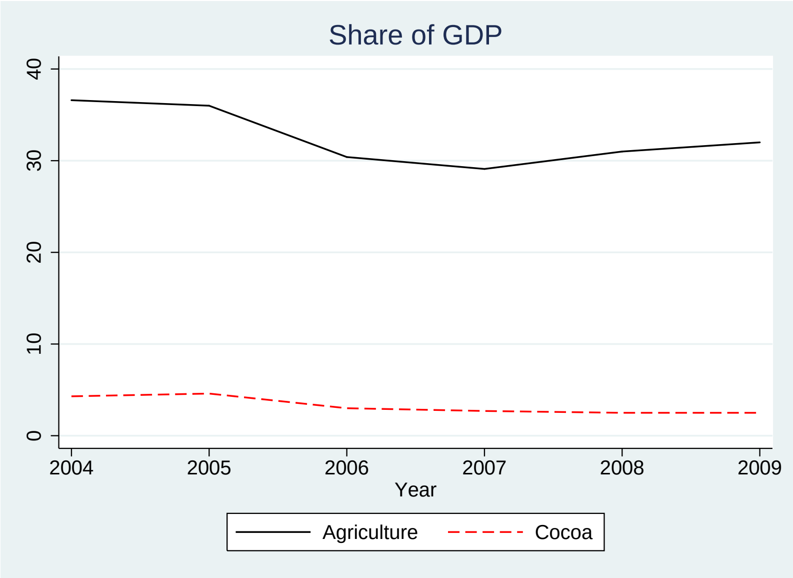
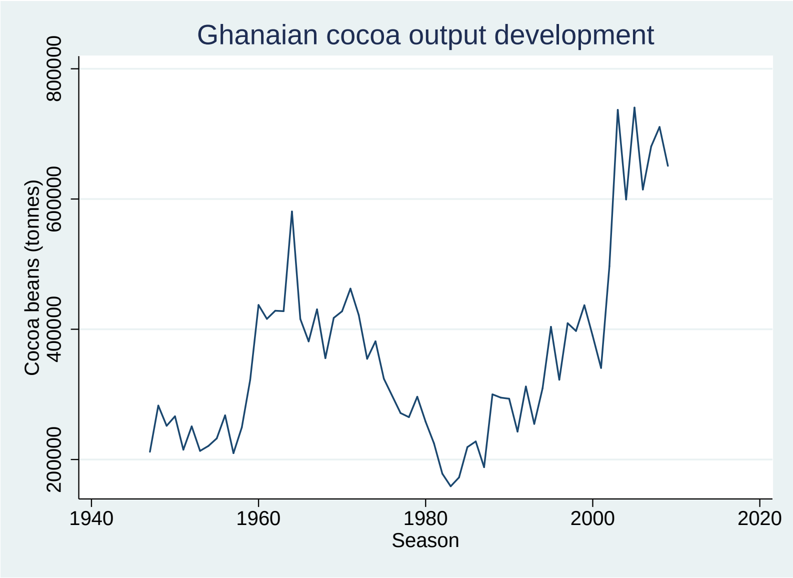
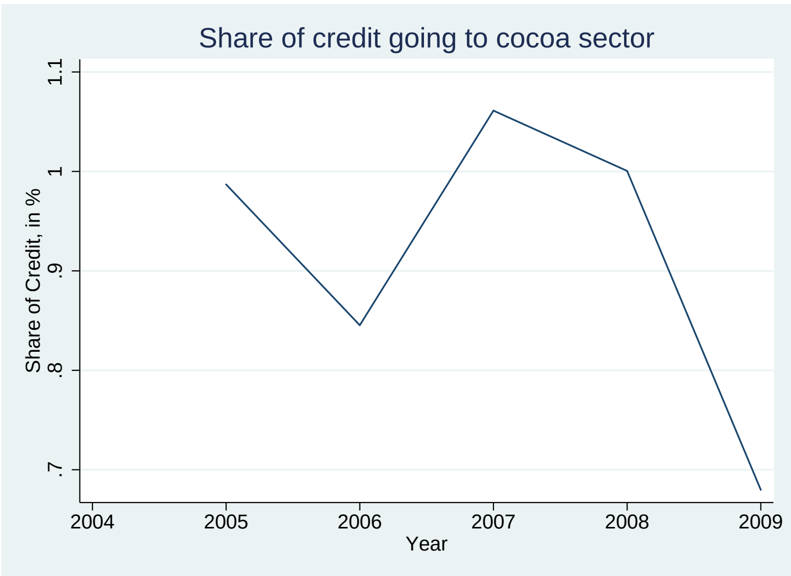


Figure 2.2: Evolution of overall cocoa production



of cocoa farmers has a direct impact on their living standards hence this issue is of interest To policy makers. The most frequent explanation for low yields of farmers is low usage of high-tech agricultural inputs such as fertilizer, insecticides and fungicides.[2] Besides other explanations of low usage of fertilizer and other inputs, the one to which a lot of attention was paid was the possible credit constraint the farmers are suffering from. In contrast to the 3% share of cocoa production and marketing in GDP, the amount of total credit going to the cocoa sector averaged around 1%. The trend in credit share of cocoa sector is shown in figure 2.3. One of the possible explanations for low amount of credit in the sector might be its riskiness, which is usually present for agricultural production on a small scale in rain-fed environment. Another possible explanation is the high correlation of risks in individual regions discouraging private sector creditors to provide credit unless covering larger regions. Similar explanations were mentioned for example in Matsumoto and Yamano (2010) and Suri (2006).

Figure 2.3: Credit share in cocoa sector



[2]Insecticides are highly important mainly due to the capsid attack threat, fungicides mainly due to the black pod disease. These are supposed to be the most threatening. The loss caused by capsid attacks has been estimated at about 25-35% per year, as mentioned by Quartey (2007). This has earlier led to the establishment of The Cocoa Diseases and Pests Control Programme in 1991.

2.2 Program description

Cocoa Abrabopa Association[3] program is a private sector initiative with the goal of improving productivity of cocoa farmers in order to increase their living standards. It was launched in 2006 by Cocoa Abrabopa Association (CAA), a subsidiary of Wienco Ghana Ltd.[4], in cooperation with NGO Technoserve. The program allocates hi-tech package of inputs designed and promoted by Cocoa Research Institute of Ghana (CRIG) since 2001 on credit on a joint liability basis. The program provides package applicable to 2 acres[5] in the first year of program participation. The package consists of 6 bags of Asaase Wura Special Cocoa Fertilizer, insecticide and fungicide package,[6] Matabi Pneumatic Sprayer, all provided by CAA. Further, CAA provides farmers with training in correct input application and good agricultural practices, which can be considered as one of the biggest advantages of the program as it (unlike many other fertilizer usage promoting initiatives) tries to minimize the loss caused by incorrect input usage. Moreover, farmers also receive business training by Technoserve, in order to make farmers treat cocoa farming as business activity.

The credit is provided to groups of 5 to 15 farmers with the joint liability. Joint liability schemes are quite popular in developing countries as they enhance monitoring and the repayment of credit. This feature is further extended by dynamic incentives. In the first year of membership, the program provides package applicable to 2 acres. If the group fails to repay the loan, then it is penalized by the exclusion from the program for the upcoming season (or more seasons). On the contrary, if the group is successful in the repayment, they might be offered by CAA promoter to extent the loan for the entire group to 4 acres for the next season (to 6 acres if they were receiving the 4 acres package). The necessity of the extension of the loan size for the whole group might be considered as one of the weaknesses of the program. To solve this issue, Zeitlin et al. (2009b) suggested to allow second year members to take additional loan of 2 acres worth on the individual liability.

The program (as well as the survey) works on the seasonal basis. The time of the first visit of CAA promoters is in January, next in February when

[3]Cocoa Abrabopa means "Cocoa for better life" in native language.

[4]Wienco Ghana Ltd. was established in 1979. It is specializing in distribution of agricultural hi-tech inputs and currently owned by Dutch and Ghanaian shareholders.

[5]2 acres present optimal area which the inputs should be applied to. Nothing prevents the farmers from using it on area of the different size, to keep it or even resell it.

[6]Insecticides and fungicides provided are: 16 bottles of 30ml Confidor 200SL, 48 sachets of Nordox and 48 sachets of Ridomil.

farmers have to make their decision about the program participation and group formation. If they decide to participate, they receive their input package in May. The harvest starts no sooner than in October and the repayment of the loans dates to December when the major part of the cocoa should be harvested.

Over the first three seasons of the CAA program, the number of members has risen dramatically: 1,440 in 2006/2007 season, 6,300 in 2007/2008 season and 11,000 in 2008/2009 season (Zeitlin et al. 2009b). Currently the number of members exceeds 19,000[7]. Contrary to the increasing number of enrolled farmers, low retention rates presented a problem (drop-out rate over 30%).

2.3 Sample description

Ghana Cocoa Board (COCOBOD) started collecting data regarding cocoa production from Ghanaian cocoa farmers back in 2002. Three waves of survey were carried out, realized every two years (2002, 2004 and 2006). For the purpose of the evaluation of CAA program, COCOBOD in collaboration with CSAE[8] expanded the scope of the survey, with annual realization (2008, 2009 and 2010 wave) allowing to evaluate the program impacts. Ghana Cocoa Farmers Survey started by the collection of data in three regions - Ashanti, Brong Ahafo, and Western - later expanded to Central and Eastern region. For the map of the regions and a chart displaying the share of regions in production see figures A.1 and A.2 respectively.

The survey contains a wide variety of information regarding farmer's plot, output, input usage, household characteristics, relationship to Licensed Buying Companies (LBCs)[9], exposure to special events and among many other also CAA membership information.[10] One of the important features for former work was the knowledge of output in particular season and of membership decision of the farmers for the subsequent one.

The distribution of age in the sample of interest (i.e. wave of survey with information about CAA) is described in figure 2.4. The mean age has been slightly increasing over the years and was around 50. The pattern did not differ significantly across members and nonmembers in a given year: While in

[7]Source: www.abrabopa.com

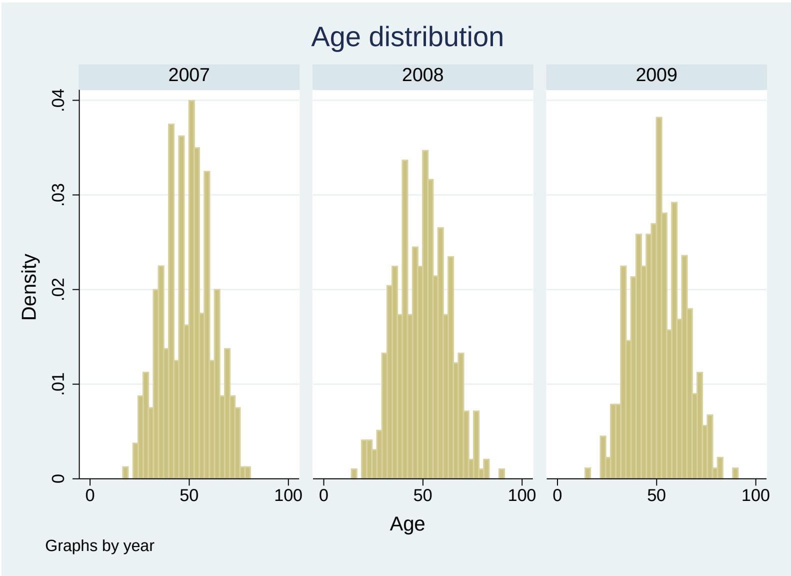
[8]CSAE = Center for the Study of African Economies, Oxford.

[9]LBCs are Ghanaian retailers.

[10]The information about the membership in CAA is covered in more recent extended version of questionnaires. For reference of older versions see http://sites.google.com/site/andrewzeitlin/data/gcfs .

2008/2009 harvest season the mean age of members was higher than in the case of non-members, the opposite was true for 2009/2010 harvest season. Summary statistics can be found in appendix (A.1).

Figure 2.4: Age distribution in the sample

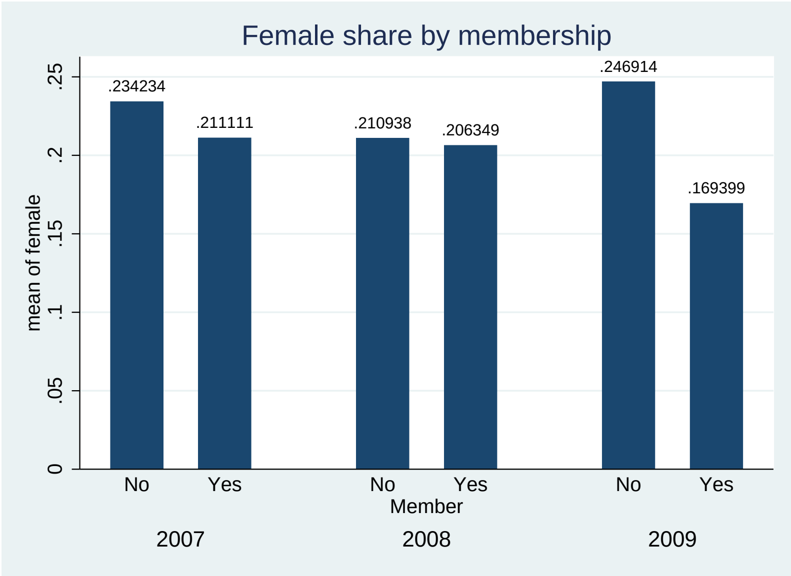


The gender distribution across farmers in our sample is not even. Moreover, women are less likely to participate in the program with the largest difference in 2009/2010 season where women form only 17% of the members of the CAA program, while presenting nearly quarter of the whole sample.[11] The evolution of female participation across the years is depicted in figure 2.5.

Farm size is another indicator largely varying for members and non-members as well as in the time dimension. The mean size of the farm remained over the seasons slightly over 4 ha while the variation across seasons was negligible. In general, this can be described as smallhold farming with negative consequences on economies of scale. Even more importantly, there is a major difference in farm size between members a non-members. This fact is depicted in figure 2.6. Huge shift in relative mean size of the farm of members and non-members signifies a major difference between those who were members in the first wave

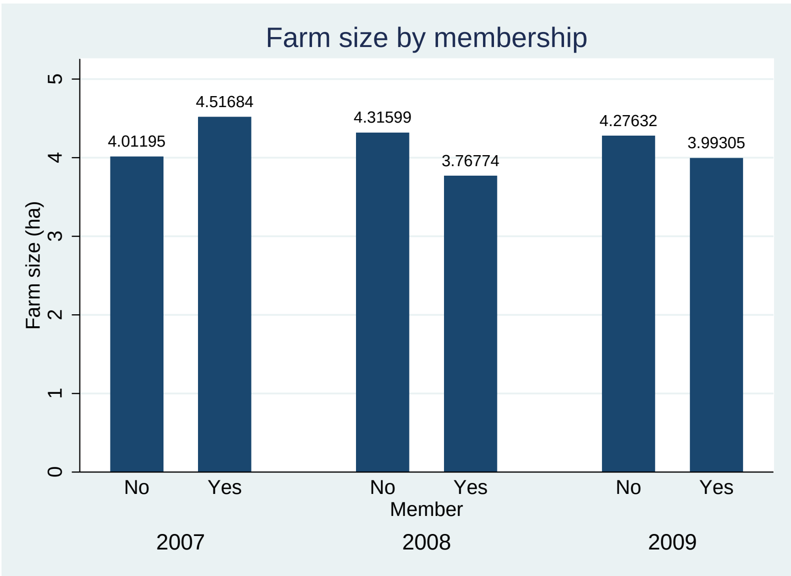
[11] Note that there is a dependence between figures in individual years as farmers have been revisited in subsequent waves of the survey. This may therefore imply either that women are less likely to participate or less likely to stay in the program.

Figure 2.5: Female share in the sample



(i.e. survey wave in 2008 collecting data about 2007/2008 harvest season) and those who have joined later.[12]

Figure 2.6: Farm size distribution



[12]Farmers staying in the program will not influence this shift unless there is a significant change in their farm size which does not seem probable in general.

In figures 2.7 and 2.8 distribution of education of non-members and members is displayed. More educated farmers are slightly more likely to join the program. This trend is most obvious in 2009/2010 season in which about 60% of non-members and more than 67% of members have JSS or higher education. More interesting are the shares of JSS or more educated farmers based on the length of their program participation - 61% for non-members, 64% for 1 year members, 66% for 2 year members and 67.5% for 3 year members. This suggests that more educated farmers are more likely to stay in the program.

Figure 2.7: Education of nonmembers

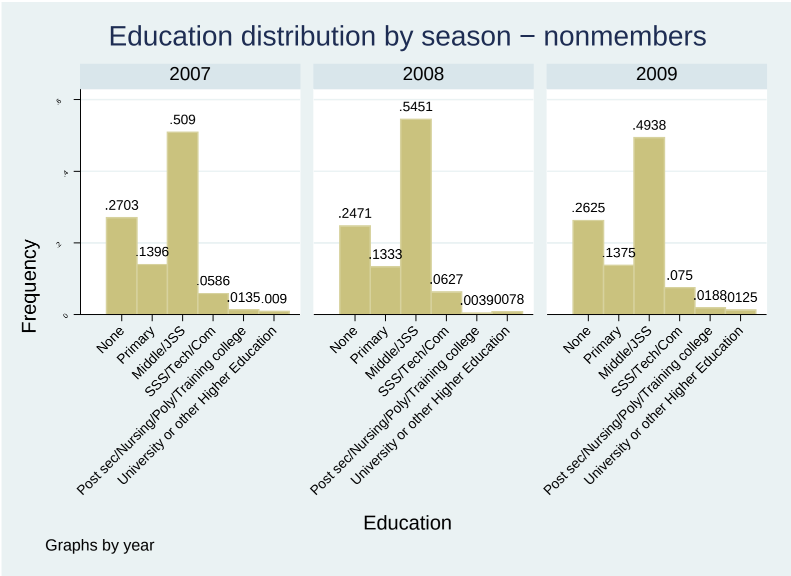
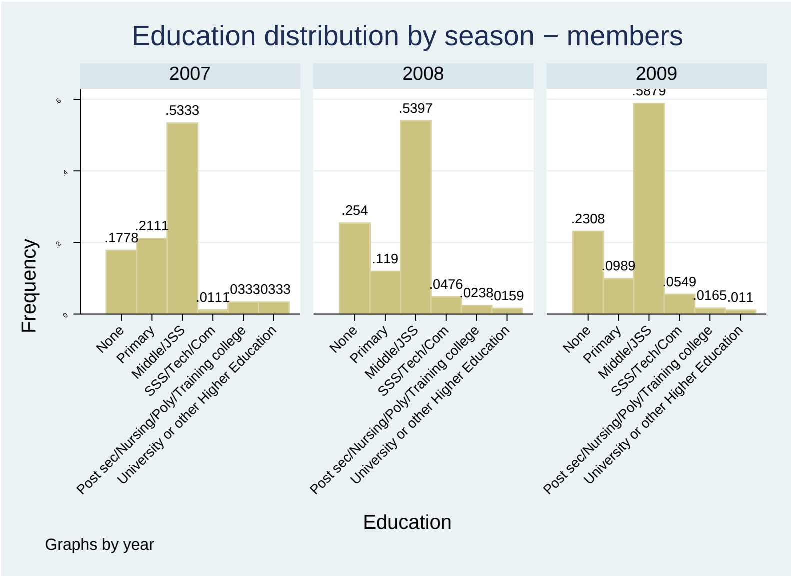
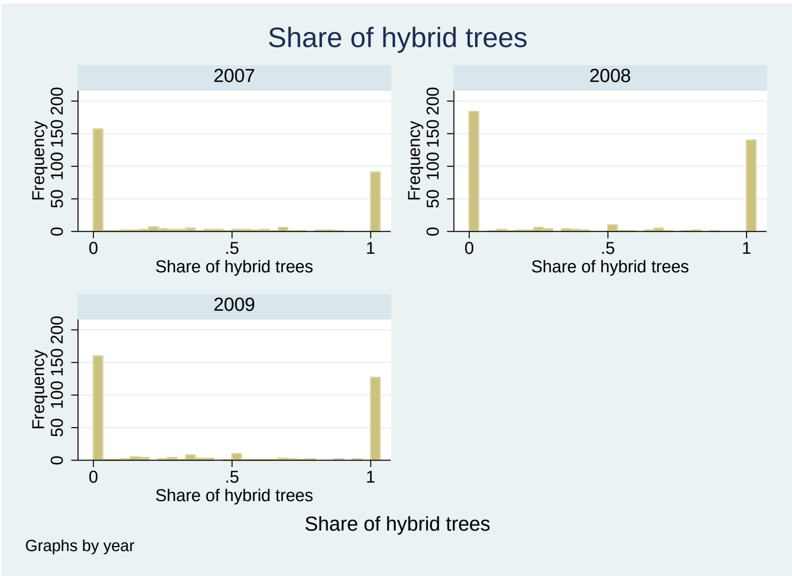


Figure 2.8: Education of members



The use of hybrid trees is limited among farmers as well, despite the fact that it has several advantages over traditional varieties - they produce more pads and start to bear fruit earlier. For reference, see Kolavalli and Vigneri (2011). Positive impact of hybrid varieties have been found in many other crops as well - maize in Malawi (Ricker-Gilbert et al. 2009) or Kenya (Suri 2006), though in the latter case substantial heterogeneity has been found as well, leaving a significant part of farmers with negative returns on adoption. As displayed in the figure 2.9, most of the farmers either fully adopt hybrid varieties or do not adopt them at all. Further, the number of farmers who do not use hybrid varieties exceeds the number of their users. The development over seasons is stable which can be partially attributed to bearing age of cocoa (3 years for hybrid varieties).

Figure 2.9: Hybrid trees usage



Chapter 3

Related work and problem statement

3.1 Literature review

The question of returns to adoption of agricultural technology is of a big interest in development economics and has been studied a lot. High mean productive impacts of technology adoption have been reported for example by Ricker-Gilbert et al. (2009), Suri (2006), Zeitlin et al. (2009a) or Matsumoto and Yamano (2010), with positive mean economic returns with the exception of the last listed work.

The question of relatively low rates of sustained use of agricultural technology despite high estimated returns presented a puzzle and as a result it has been studied a lot in the literature. Several possible explanations have been found. Zerfu and Larson (2010) found on the sample of Ethiopian farmers that constraints in credit market, high transport costs and illiteracy have a negative impact on the adoption of agricultural technology. Another explanation was suggested by Duflo et al. (2009) in a study on Kenya. Their explanation is of a behavioral nature - according to them the farmers who are patient (and therefore willing to save funds in order to purchase fertilizer) tend to underestimate the probability that they will turn impatient in the forthcoming periods. In other words, they postpone adoption of agricultural technology as they expect themselves to adopt it in the future but they are more likely to spend their funds in a different way than they think they would. As a result, time limited 17 % discount on fertilizer in the time of harvest had the same impact on the amount of purchased fertilizer as 50 % discount with free delivery in the time of fertilizer application.

An alternative explanation of low adoption rates of agricultural technology is the heterogeneity in returns. It stems from the limited amount of information provided by average treatment effect as it is possible that individual returns vary heavily across farmers. Heterogeneity in returns has been studied by Duflo et al. (2008), though evidence for heterogeneity was lacking. On the contrary, Suri (2006) found substantial heterogeneity in the study on usage of hybrid varieties by Kenyan maize farmers. She found the distribution of farmers returns and pointed out that a small fraction of farmers who would benefit the most from adoption did not adopt (due to limited input access) and contrarily large fraction of adopters have individual returns close to zero.

Similar practice has been conducted by Zeitlin et al. (2010) on the sample of Ghanaian cocoa farmers (used in this thesis) to estimate quantile treatment effect of the first year of CAA program participation. Zero return on program participation could have not been rejected for the lower quartile of the population (by individual returns). This paper further contributed by enhancing the model for program retention. The authors have found that not only the repayment failure of farmers themselves but also the repayment failure of their group members has a statistically significant and negative impact on the likelihood of retention. Moreover they found a positive relationship between individual return on program participation and the probability of staying in the program.[1]

Another reason for low rate of adoption provided in the literature is possible large importance of social learning. In such case farmers' decision on program participation would heavily rely on their formation of the informational network and on their experience with input levels[2] as well as that of their information neighbors. In a study on Ghanaian pineapple farmers Conley and Udry (2005) have shown that social learning plays an important role in diffusion of new technology and that the effect is the highest for inexperienced farmers. Zeitlin (2010) used the sample of Ghanaian cocoa farmers and confirmed the importance of the role of social learning in technology adoption using

[1] They used two period change in cocoa as a proxy for true realized individual return. Further they have found high correlation between this measure of realized return and quantile treatment effect - alternative measure for individual return.

[2] In case of CAA the social learning will cover information about adoption of program as well.

farmers visiting the same LBC (Licensed Buying Company) as an information network and allowing for endogenous group formation.

Duflo et al. (2008) studied Kenyan maize farmers and pointed out that the authorities' recommendation of input usage based on the results from experimental fields might be in fact not profitable for the vast majority of farmers who are, unlike researchers on experimental plots, exposed to real world problems and are unable to acquire complementary inputs or they do not apply the inputs correctly. They found that for the majority of farmers the use of optimal quantity of fertilizer increases return with a little risk added. This might be the case of CAA program, because it promotes and provides packages recommended for 2 (4, 6) acres designed by CRIG. This problem is further augmented by the design of micro-credit scheme requiring the same amount of fertilizer provided to each farmer in a group.

The papers written by Zeitlin et al. (2009a), Zeitlin et al. (2009b) have studied the case of Ghanaian cocoa farmers as well and can serve as a good starting points for this thesis. These papers presented the basic identification strategy for the evaluation of agronomic and economic impacts of the program using the pipeline method. They concluded that average treatment effect among farmers is high and from this point of view it seems as a great success. They also pointed out that despite the quite high repayment rate of the loans, large share of the program members choose to leave the program. Their inability to repay is not sufficient to explain this high drop-out rate. To investigate this question, they estimated probit model with dependent variable program retention and reached the conclusion that not only repayment problems of the farmers have a negative impact on retention but also low individual returns on program makes farmers more likely to leave.[3] As noted above, in Zeitlin et al. (2010) this model have been extended by an additional control variable for peer repayment failure.

3.2 Problem statement

The early works measuring average treatment effects of CAA program - Zeitlin et al. (2009a), Zeitlin et al. (2009b) - have estimated significant positive ATE. More importantly, the calculated economic return on the program participation was positive. What presented a puzzle was the fact that many farmers stepped

[3] As a proxy for individual return on program change in ln cocoa has been used.

out of the program even when they did not have problem with the repayment of the loan. This contrasted estimated positive returns. Nevertheless high estimated impact has advocated widespread adoption of the program.

Several explanations of why farmers do not join the program or why they leave were provided in section 3.1. Namely Zeitlin et al. (2010) have focused on the distribution of returns. By estimating quantile treatment effect they have found evidence of substantial heterogeneity in return on program participation among farmers. They were unable to reject the hypothesis of zero economic impact for lower quartile of farmers. Why are some farmers unable to generate positive economic return with a use of hi-tech inputs was not clear. It could be caused by their permanent inability to do so or by a transient shock. In either case, this has put a shadow of doubt on the promotion of the widespread adoption of the program. If some farmers are unable to make the use of hi-tech inputs profitable, then it makes no sense to promote adoption for all farmers.

Studies mentioned above focus on the estimation of the impact on a sample of 1 year members (along with non-members). This is one of the limitations of these studies because the returns to program participation are expected to be higher in the later years of sustained membership. This problem has not yet been studied. If there is any major difference in productive and especially economic returns on the program based on the length of farmers participation in the program, then it would have important policy consequences. There are few reasons why the agronomic and economic returns in later years are expected to exceed the returns in the first year of participation. First, the increase in the return (in absolute value) is rational as farmers have the opportunity to extend the size of the loan in the later years of membership. From the similar rationale stems the expected higher increase in yields because farmers with larger plots are able to cover a larger share of their plot by package provided by CAA promoters. Second, nutrient stock in the soil on cocoa farms depletes without the use of inorganic fertilizer. On fertilized plots the addition of nutrients is likely to exceed the loss caused by growing of cocoa. For reference, see Hartemink (2005). As a consequence, longer fertilizer usage has higher cumulative impact on nutrient stock in a soil hence sustained usage of fertilizer is expected to have increasing productive (and economic) impact.

The objective of this thesis is the estimation of long-term productive and economic impacts of sustained participation in the CAA program. As mentioned above, return to program participation is expected differ with respect to length of membership. However, this issue has not been studied in the earlier literature. The hypothesis is that both are increasing in the length of participation in the program. If so, this has important policy consequences. First, the drop-out of farmers presents a more painful issue as it prevents farmers from reaching those higher returns. Second, some of the farmers realizing negative return in the first year of program participation might reach positive return if they participated for more years. Chapter 4 provides a model for the estimation of productive impacts, chapter 5 estimates the economic impacts of longer program participation. Further, in this thesis we want to confirm results found by Zeitlin et al. (2009b), Zeitlin et al. (2010) regarding the program retention applying similar approach on the sample of all farmers (not only 1st year members) - that there is a positive relationship between farmer's individual return on the program and program retention and negative one between a repayment problem (both of farmer and other group members) and retention. To do so, we have to adjust the model of program retention because there is a difference in the amount of experience with the returns on program participation resulting from different length of it.

Chapter 4

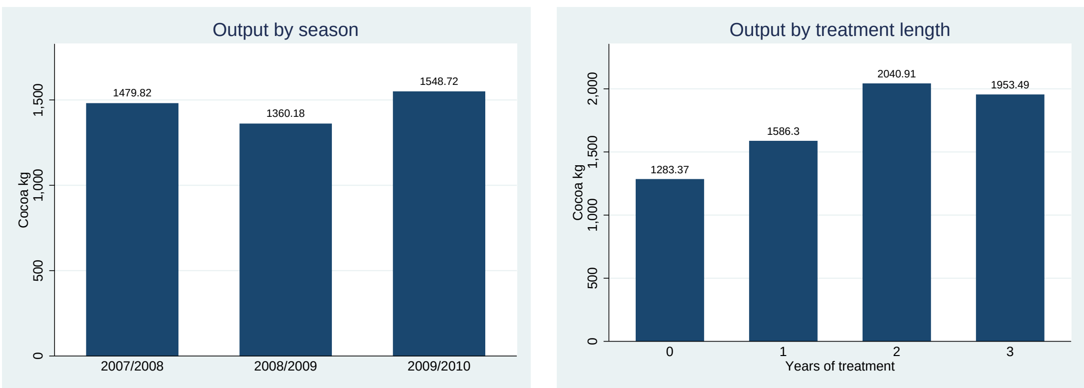
Evaluation of the impact of the program on output

4.1 Output overview

The total Ghanaian production of cocoa beans over the last five recorded years was between 600,000 and 750,000 tonnes per year. The average national productivity remains near 400 kg/ha, which is below the productivity of other countries producing cocoa beans. [1]

In figures 4.1 average cocoa output in kg in individual growing seasons and for groups with same length of treatment is displayed.

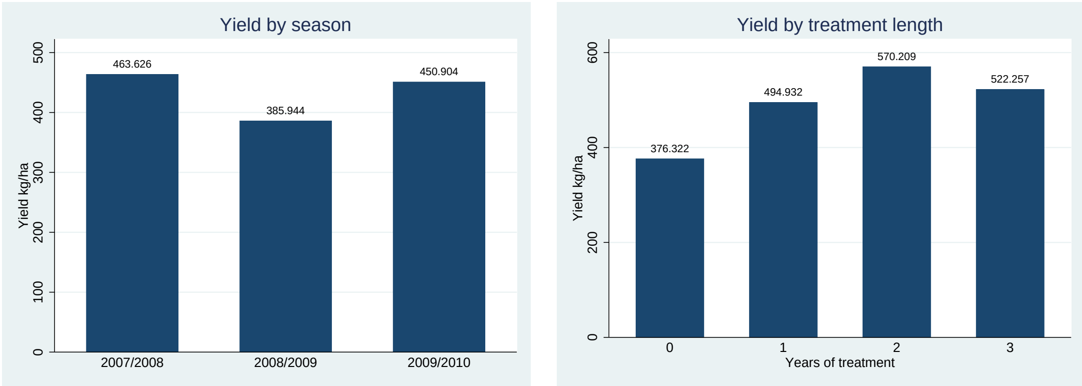
Figure 4.1: Output by seasons and treatment



On the same basis yield in figures 4.2 is displayed.

[1] Exact figures are displayed in table A.2

Figure 4.2: Yield by seasons and treatment



As we can see from the figures, in growing season 2008/2009 there was a significant drop in both yield and output possibly caused by negative weather conditions. The identification of the causes of this drop is not within the scope of this text. The similarity of the seasonal trends of output and yield signifies that possible changes of farm size were not too large to play a major role in possible changes in yield of farmers. In the case of length of the treatment, higher output and yield for members was expected. The decrease of both output and yield between 2 and 3 year treatment therefore seems a bit puzzling. This issue and overall identification of impacts will be addressed in the following sections.

4.2 Basic model specification

Basic model used for the estimation of productive impacts of the program will be the production function of a cocoa in a reduced form (as in Zeitlin et al. (2009a))

$$Y_{it} = \boldsymbol{\beta M}_{it} + \boldsymbol{\gamma Z_{it}} + \mu_i + \nu_{vt} + u_{it} \tag{4.1}$$

where $\boldsymbol{M_{it}}$ is a vector of membership characteristics, $\boldsymbol{Z_{it}}$ is a vector of other control variables, μ_i represents individual characteristics of a farmer, ν_{vt} is village-specific effect of village v in year t and u_{it} is idiosyncratic error of farmer i in time t.

As already described in Zeitlin et al. (2009b), Zeitlin et al. (2009a), simple regression of cocoa output on membership status will not provide information about the impact of the program. They mentioned several possible biases which would occur if a such naive approach was used. First, what if, on average, more productive villages were visited by CAA promoters for the first time in earlier years and less productive later (i.e. $E(\nu_{vt}M_{it}) \neq 0$) ? Second, what if the more productive farmers are more likely to join in the program (i.e. $E(\mu_i M_{it}) \neq 0$)? In such case, simple regression of output on membership status would overestimate program's impacts. And what if there is a systematic difference between those who join the program in the first year of presence of CAA promoters in their village and those who decide to join later (after observing their neighbor's outcomes)? All those effects might be a source of possible bias in a naive estimation procedure.

Zeitlin et al. (2009b), Zeitlin et al. (2009a) suggested two possible strategies to identify the first-year impacts of the program relying on different sets of assumptions. The key feature of the dataset for both strategies was the ability to collect information about the future membership of farmers and therefore not only about output of current (and already treated) members, but also about output of future members (who joined the program but have not been treated yet). The starting point was the division of farmers into 4 groups:

1. Recent members - decided to join and were treated
2. Recent nonmembers - decided not to join hence were not treated
3. Future members - decided to join but not been treated yet
4. Future nonmembers - decided not to join for upcoming season (were not treated)

First of the strategies proposed to compare output of recent members with output of future members and estimate the average treatment effect in a following way:

$$ATE_1 = \text{Output(Recent member)} - \text{Output(Future member)} \tag{4.2}$$

Alternatively, using the notation based on the equation 4.1

$$ATE_1 = E(y_{it}|T_{it} = 1, M_{it} = 1, Z_{it}) - E(y_{it}|T_{it} = 0, M_{it} = 1, Z_{it}) \tag{4.3}$$

where T_{it} is a dummy variable denoting whether the farmer had option to be treated in the current year and M_{it} a dummy for membership status. Such

comparison will deal with possible self-selection bias caused by systematic difference between those who decide to join and those who decide not to join. The problem is that this strategy relies on the assumption of random village selection (i.e. $E(\nu_{vt} M_{it}) = 0$).
The second identification strategy they proposed deals with the problem of within village selection bias in a different way. It uses a Difference-in-difference estimator in a following form:

$$\begin{aligned} ATE_2 =& [\text{Output(Recent member)} - \text{Output(Recent nonmember)}] - \\ & [\text{Output(Future member)} - \text{Output(Future nonmember)}] \end{aligned} \tag{4.4}$$

Again, using alternative notation

$$\begin{aligned} ATE_2 =& [E(y_{it}|T_{it} = 1, M_{it} = 1, Z_{it}) - E(y_{it}|T_{it} = 1, M_{it} = 0, Z_{it})] - \\ & [E(y_{it}|T_{it} = 0, M_{it} = 1, Z_{it}) - E(y_{it}|T_{it} = 0, M_{it} = 0, Z_{it})] \end{aligned} \tag{4.5}$$

The first term in equation 4.4(4.5) is equal to the difference between output of members and nonmembers of the program in the villages where the treatment was available and consists of both Membership Effect and Treatment Effect, while the second term compares the output of future members and nonmembers and captures the Membership Effect. Of a key importance is the usage of both members and nonmembers which allows to account for village-fixed effects using fixed effects estimator. This strategy relies on the assumption of no externalities of the program, however its violation will, if the externalities are positive, lead to underestimation of the impact of the program.

Assumption 4.1 (No externalities). The program has no externalities on non-members.

$$E(\mu_{it}|M_{it} = 0, Z_{it}) = E(\mu_{it}|T_{it}, M_{it} = 0, Z_{it})$$

4.3 Extended model specification

The second method from the previous section is a good starting point for our extended identification strategy used to obtain the estimate of impact of 1st, 2nd and 3rd year of continuous membership on the output. This extension was not provided in the literature, though is necessary for the estimation of 2nd and 3rd year impacts. Even in this case the idea is the same - the evaluation of the total effect of membership in individual years and consequent subtraction of membership effect to obtain estimates of average treatment effect.

The problems faced in the case of identification of 1st year impact are even more noticeable in the extended version used for the estimation of the impacts of longer membership. The questions which are needed to be taken into account are of a similar nature - what if those who are willing to join the program are more productive than those who are not? What if more productive farmers are willing to stay in the program while their less productive counterparts leave? Zeitlin et al. (2010) and Zeitlin et al. (2009b) have estimated positive relationship between return on the program and retention rate. Furthermore, the estimated impact of membership on input demand has shown that the change in input usage does not substantially affect average return. This hints that there are considerable differences between those willing to stay in the program and those willing to leave (as higher output cannot be considered to be a result of higher input usage).
In order to estimate the effect of long-term membership, the estimation procedure requires to use the data collected in more than one year (2008, 2009 and 2010 survey rounds). As a consequence, it becomes necessary to consider time-varying (climatic etc.) conditions affecting output, perhaps with varying impact across locations.
This leads to an important assumption required for the consistency of the estimation:

Assumption 4.2 (Stable membership effect). The membership effect is the same regardless of the time of the first visit of a village by CAA promoters.

$$E(\mu_{it}^{Memk}|M_{it}^{kyrs} = 1, T_{vt} = 1, Z_{it}) = E(\mu_{it}^{Memk}|M_{it}^{kyrs} = 1, T_{ur} = 1, Z_{it}) \forall u, v; \forall t, r$$

Having mentioned the most severe pitfalls and being armed with the crucial assumptions, it is time to introduce the set of membership controls and the strategy for identification of impacts.
The sample of farmers is, as in the case of estimation of 1 year impact, divided into groups based on the length of treatment of the farmers and their membership decision for the upcoming season.[2] This basic identification strategy leads

[2]It is important to note that we are interested in continuous membership, hence the membership controls are accounting only for such stable membership behavior.

to 8 reference groups summarized in the following table:

Group	**Group name**	**Yrs of treatment**	**Future membership**
1	Non-members	0	No
2	Future 1y members	0	Yes
3	Non-future 1y members	1	No
4	Future 2y members	1	Yes
5	Non-future 2y members	2	No
6	Future 3y members	2	Yes
7	Non-future 3y members	3	No
8	Future 4y members	3	Yes

The division of the farmers into the above listed groups might lead to few caveats. First, there might be a significant difference between output of those willing to join in the first year of the visit of their village by CAA promoters. Moreover, there might be even different return on the program based on the timing of joining the program.

Second, what if there is a significant impact not only of a continuous participation, but also of interrupted participation. In other words - is it possible to treat in the same way those who were treated for one year (two years, . . .) and those who were treated for more seasons but not in the last one?

Third, are not the first two groups (group 1 and group 2) overly heterogenous in a sense of their past participation conditions? It is possible that some of the farmers benefited from their participation but left the program due to negative experience with a group liability scheme.

To address all the caveats mentioned above, control variables are introduced for the timing of joining of the program, rejoining farmers and for past members. Still, the first caveat is addressed only partially. We only use controls for difference in outputs. The consistency of the estimator however relies on the following assumption:

Assumption 4.3 (Return homogeneity in timing of joining). The membership effect is the same regardless of the time of entering the program.

$$E(y_{it}|M_{it}^{kyears}, Z_{it}) = E(y_{it}|M_{it}^{kyears}, Z_{it}, \text{Timing}_i)$$

The last group of variables included in the model is the set of farmer/plot characteristics. The motivation for their inclusion is to cope with possible omitted variable biases (at least the most severe ones). First of the farmer characteristics included is gender. Quite possibly men and women have different output (through gaps in yield) and more importantly gender is expected to have correlation with program participation. Second farmer control is education - dummy variable denoting whether the farmer has JSS or higher education - included for the same reason. Further - two plot specific variables have been added: Plot size covered by hybrid trees and the size of the plot not covered by hybrid trees. This resolves the possible correlation between plot size and the membership. If farmers with bigger plots are more likely to participate, then the estimate will suffer from omitted variable bias. Distinction between hybrid land and non-hybrid land is motivated by possible differences in their yields. Similar foundations were for example in the case of Kenyan maize farmers (Suri 2006). Moreover their inclusion advocates the level form of the model due to the elimination of the impact of farm size. As the program provides credit on fixed amounts of inputs (2, 4 & 6 acre packages), the percentage change of output would largely differ across farmers (Larger percentage change will occur in the case of farmers with smaller plots).

Having constructed the model, one important question remains unanswered - how to identify the impacts? The strategy is similar to the one used in a simple model of 1 year impact estimation, though with few complications. The model does not directly provide the estimates of total impact of 1 (2, 3) year membership (i.e. treatment and membership effect). Nevertheless, it is possible to proceed in the following way:

$$\begin{aligned} ATE_{1y} &= \underbrace{E(Y_{it}|Z_{it}, \text{1y tr.})}_{\text{Total effect}} - \underbrace{E(Y_{it}|Z_{it}, \text{Future 1y member})}_{\text{Membership effect}} \\ &= E(Y_{it}|Z_{it}, \underbrace{\text{1y tr., Future member}}_{\text{Future 2y member}})\, \underbrace{P(\text{Future member}|\text{1y tr.})}_{p_{f2y}} \\ &+ E(Y_{it}|Z_{it}, \underbrace{\text{1y tr., Future non-member}}_{\text{non-future 1y member}})\, \underbrace{P(\text{Future non-member}|\text{1y tr.})}_{p_{nf1y}} \\ &- E(Y_{it}|Z_{it}, \text{Future 1y member}) \end{aligned} \tag{4.6}$$

The output expectations are estimated by the model, so the only unknown in the expression above is the probability of staying in the program after one year of treatment. However it is still possible to estimate it in the following way:

$$\hat{p_{f2y}} = \frac{n(\text{Future 2y member})}{n(\text{Future 2y member}) + n(\text{non-future 1y member})} \tag{4.7}$$

$$\hat{p_{nf1y}} = 1 - \hat{p_{f2y}} \tag{4.8}$$

Along similar lines will be derived the average treatment effect of 2nd and 3rd year of program participation, though slight adjustment (or rather remark) is necessary. When subtracting future 2 year (3 year) members to clean out the membership effect, the average treatment effect of 1 (2) years is subtracted as well, so it has to be added back to identify the effect of longer participation in comparison to untreated farmers.

4.4 Estimation and results

Previous sections were dedicated to dealing with the problems caused by individual characteristics. To address potential differences between seasonal agronomic conditions (climatic etc.) and village specific characteristics the fixed effect estimator with village-year fixed effect (for each combination of year and village) is used. The estimates of the model are presented in the table A.3 and are reported with clustered standard errors. The average treatment effects based on the identification strategy described in previous section are summarized in the table 4.1:[3]

Avg. treatment effect					95% Confidence Interval	
Cocoa kg	Coef.	Std. err.	t	p-value	lower	upper
1st year	396.7579	151.5006	2.62	0.011	95.01832	698.4975
2nd year	689.0957	224.531	3.07	0.003	241.9034	1136.288
3rd year	661.318	191.02	3.46	0.001	280.8686	1041.767

Table 4.1: Output average treatment effect estimates

The estimated productive impacts are significant on 5 % level of confidence and unsurprisingly positive for all the studied lengths of participation. With

[3]Note that the standard errors, p-values and confidence intervals are not 100% correct as they are not adjusted by variance of the estimator 4.7. Nevertheless, due to the size of the sample the variance of such estimator is extremely low and therefore its impact on the precision of the estimated ATEs is negligible ($\widehat{Var(\hat{p})} = \frac{\hat{p}\hat{q}}{n}$). The same consideration will be applied in all cases of this thesis.

the mean output of untreated farmers equal to 1283 kg, the estimated impacts represent 31%, 53% and 51.5% increase in output for 1, 2 and 3 year participation respectively. Nevertheless, the results raise a few questions. First, the average treatment effect for the 1 year participation in the program is lower than in the related work. Zeitlin et al. (2010) have used a set of estimators, while the fixed effect estimator with observable farmer controls predicted the average treatment effect of 527.4 kg. Even higher estimated productive impact of 742 kg was estimated by Zeitlin et al. (2009a), though the model did not use any control variables. One of the possible explanations is the omitted variable bias. Another possibility is that this work uses larger sample, while newly included farmers might have lower marginal response to fertilizer (e.g. due to different soil types). [4] Varying response to fertilizer in different areas was studied for example by Graybeal (2000).

Second, the rise in productivity in the second year of membership seems to support our hypothesis (as long as there is no substantial change in input usage). On the contrary, a small decline in cocoa output has been found in the third year of program participation (though difference between 2nd a 3rd year impact being equal to zero cannot be rejected). It is possible that it might be caused by measurement errors mainly in output and/or size of the farm. This decrease is not present if we restrict our model to sub-sample leaving outliers out of the model.[5] In such case there is no longer decrease in output for 3rd year. Alternative explanation can be that the stable membership effect assumption (assumption 4.2) is violated - that future 3 year members might be more productive than present 3 year members.[6] Another possible explanation for no increase between 2nd and 3rd year of program participation (contrary to the case of big increase between 1st and 2nd year) might be that there was a higher demand for the increase of the loan size (to 4 acres worth) in the second year of participation, but it remained stable in the third year. Other explanation that comes to mind is that 3rd year members (in a sense of treatment)

[4] This might be true even though those farmers are not less productive generally, only they are different in sense of response to fertilizer.

[5] As an outlier farmer with yield outside interval (30;2000) kg/ha is considered. Yield seems reasonable measure for capturing possible measurement errors (unlike output which is heavily dependent on the size of the farm). ATE for this restricted sample is attached in table A.5

[6] In the case that the future 3rd year members who were visited in 2008 for the first time have higher individual returns on program, extremely high 2nd year returns (unlike 3rd year) might be found. This would result from the fact that their 3rd year output - expected to be extraordinarily high as well - is not recorded in the sample.

might have limited access to complementary inputs resulting in suboptimal production. The last two possibilities will be addressed in the subsequent sections.

Third, the assumption 4.3 is still strong. The heterogeneity in returns correlated with the timing of joining the program might be a source of bias. Still the model can be estimated when we restrict ourselves on early adopters to estimate the average treatment effect for this subpopulation. The same practice cannot be conducted in the case of later entry to the program due to the small size of such a subpopulation. This estimate of impact on early adopters results in higher treatment effect (with the exception of 3rd year member where the estimate is slightly lower). For exact values see table A.4. Our preferred estimate is the original one as it covers also farmers less suitable for adoption and is therefore more conservative and representative as average treatment effect.

Fourth, the estimated results might be influenced by the fact that the size of the sample is quite limited. This regards mainly the 3rd year members whose count is about 40 in our sample. As a result, this might present a limitation to achieving the asymptotic properties of the estimator.

Finally, the agronomic return on received inputs has to be calculated by dividing the estimated average treatment effects by average size of the package received (measured as number of acres to which it is recommended). The average dose received is 2.2, 2.7 and 3.5 acres for group of treated for 1, 2 and 3 years respectively.[7] This results in 180 kg/acre, 255 kg/acre and 189 kg/acre in first, second and third year of program participation respectively. As in the case of average treatment effects, the agronomic return in the 2nd year does not only exceed the return in the 1st year of program participation, but also return in the 3rd. Possible explanations are similar as for average treatment effect. The difference between 1st and 3rd year agronomic impact does not seem large enough to support the hypothesis of growing agronomic returns on program. This could be attributed to possible decline in market purchases of inputs. This issue will be covered in chapter 5.

[7]These figures might not be exact as a substantial part of the dataset regarding the size of dose received was incomplete or corrupt. In order to provide applicable outcome the data has been cleaned. The cleaning procedure is described in the appendix.

Chapter 5

Evaluation of economic impacts

5.1 Cost-benefit mechanism

As already mentioned by Zeitlin et al. (2009b), Zeitlin et al. (2010), having estimated the productive impacts of the CAA program does not allow us to calculate correctly the economic returns on the program. A simple division of the estimated productive impact by the cost of the investment would not be sufficient to fully assess the economic impacts of the program. As they pointed out, it is necessary to take into consideration possible changes in the demand for inputs resulting from program participation.

The explanation why is the demand for inputs expected to be affected by the program participation is straightforward (and more importantly, even testable). Due to the fact that the program provides hi-tech package of inputs additional units of labor are required for their application. From this point of view, it does not seem very probable that the set of inputs demanded by a farmer will change only in the subset of those provided by the CAA program due to their complementarity with labor inputs. The second channel of a possible shift in demand for inputs is an increase in output[1] resulting in higher labor requirements during the harvest. While the hi-tech inputs provided by CAA can be considered to be complements of labor inputs, they are substitutes of market-purchased fertilizer, insecticides or fungicides. If CAA program is successful at the alleviation of farmers' credit constraint then it is reasonable to expect a change in factor demand.

To compute the average economic impacts of the CAA program, we have to account for a possible shift in factor demand, otherwise the economic impact

[1]Productive impacts has been estimated in chapter 4.

might be driven by an increase in input usage rather than by program itself. The economic impacts will be calculated in the same way as in Zeitlin et al. (2009b), Zeitlin et al. (2010):

$$ROR_{k-years} = \frac{p_y ATE_{k-years} - p_{\boldsymbol{X}} \Delta \boldsymbol{X}_{k-years}}{ACI_{k-years}} - 1 \tag{5.1}$$

$AER_{k-years}$ denotes average economic return on k years of program participation, p_y average price of cocoa, $p_{\boldsymbol{X}}$ vector of average prices of inputs[2], $\Delta \boldsymbol{X}_{k-years}$ change in demand for factors due to the program participation and finally, $ACI_{k-years}$ represents the average cost of investment for group of k-years participants. The equation 5.1 still contains one unknown - the vector of factor demand shifts. Next two sections are dedicated to their estimation.

5.2 Non-labor inputs

While there are various non-labor inputs used in cocoa production, we focus on the most important factors which are provided by CAA program in hi-tech package. These are fertilizers, fungicides and insecticides. First thing of interest is whether do farmers use the above mentioned inputs at all. This issue is summarized in table 5.1.

	Non-members			Members		
Season	fertilizer	fungicides	insecticides	fertilizer	fungicides	insecticides
2007/2008	.285	.428	.910	.989	.804	1
2008/2009	.416	.523	.760	.992	.906	.984
2009/2010	.528	.654	.926	.995	.882	.984

Table 5.1: Input usage - fraction of farmers

The fact that the input usage in all three cases amongst members dominates the usage amongst non-members is not surprising. Nearly 100% usage of them amongst CAA members more or less confirms that they are not (at least on a large scale) keeping or reselling the input package provided by CAA promoters. What was especially alarming was only 28.5% usage of fertilizer amongst non-member in 2007/2008 season. Fortunately, this figure has increased during the subsequent seasons. This might have basically three explanations. First, it is possible that a part of the farmers who were not using these inputs at all

[2] Average here is meant in a sense of average across farmers.

joined the program and then started using them. Second, it is possible that they started using these inputs after observing the outcomes of their peers. If so, this can be treated as a positive externality of the CAA program. In such case the estimation of impacts would underestimate the social benefit of the program. Third, this pattern could signify some village selection problem.

The figures of input usage in absolute values are displayed in tables 5.2 and 5.3. Again, quantities applied by members dominate the quantities applied by non-members. As obvious from tables 5.1 and 5.2[3], in seasons 2007/2008 and 2008/2009 median use of fertilizer amongst non-members was still zero. Nevertheless, the mean usage of fertilizer has substantially risen even in the case of non-members, again suggesting the presence of externalities. While in 2007/2008 season members used on average approximately 3.75 times more fertilizer than non-members, in 2009/2010 this ratio was only about 1.5.

	Non-members					
	Fertilizer 50 kg bags		Fungicides Sachets*		Insecticides Litres*	
Season	mean	median	mean	median	mean	median
2007/2008	2.51	0	79.09	50	4.97	4
2008/2009	4.83	0	73.51	47	6.65	3
2009/2010	8.74	4	66	40	6.13	5

Table 5.2: Average and median input quantities - non-members

	Members					
	Fertilizer 50 kg bags		Fungicides Sachets*		Insecticides Litres*	
Season	mean	median	mean	median	mean	median
2007/2008	9.37	6	87.78	96	10.80	6.8
2008/2009	11.46	6	106.77	96	22.49	16
2009/2010	13.25	11.5	94.45	96	19.06	16

Table 5.3: Average and median input quantities - members

To identify the impact of CAA program on factor demand, we introduce a simple factor-demand equation in a similar manner as in case of output

[3]Table 5.2 contains figures from sample of farmers who filled the quantities in litres and sachets for insecticides and fungicides respectively. The observations with other units of measurement were not used for computation, because the conversion rates are not known to the author of the thesis. Later on, this pitfall will be highlighted.

(equation 4.1). The same approach was used in Zeitlin et al. (2009b), Zeitlin et al. (2010). Our factor demand equation has the following form:

$$\boldsymbol{X}_{it} = \boldsymbol{\delta M}_{it} + \boldsymbol{\phi Z_{it}} + \zeta_i + \theta_{vt} + \epsilon_{it} \tag{5.2}$$

The identification strategy is the same as the one introduced in chapter 4. The same is true for the estimation procedure and the set of assumptions required for consistency of the estimator. However, there is an additional assumption resulting from data structure in case of insecticide and fungicide demand:

Assumption 5.1 (Randomness of measurement unit). There is no systematic difference across groups of farmers using different insecticide/fungicide measurement units.

Under our set of assumptions the estimator will be consistent. The estimation results are provided in table 5.4.

Avg. Treatment Effect					95% Confidence Interval	
Fertilizer, 50 kg bags	Coef.	Std. err.	t	p-value	lower	upper
1st year	3.582714	1.503842	2.38	0.020	.5875525	6.577875
2nd year	3.77979	1.901313	1.99	0.050	-.007004	7.566584
3rd year	4.12198	1.567246	2.63	0.010	1.000538	7.243421
Insecticide, litres						
1st year	11.00257	3.122129	3.52	0.001	4.784312	17.22083
2nd year	10.89682	6.150418	1.77	0.080	-1.352801	23.14644
3rd year	12.20896	3.90175	3.13	0.002	4.437948	19.97997
Fungicide, sachets						
1st year	25.677	9.801637	2.62	0.011	6.155351	45.19866
2nd year	37.47602	10.26427	3.65	0.000	17.03295	57.91909
3rd year	22.50802	16.3946	1.37	0.174	-10.14467	55.1607

Table 5.4: Demand estimates - non-labor inputs

For all three productive factors the program leads to increase in their usage. Nevertheless, the quantities by which the demand increases are lower then the quantities allocated by the program[4] hence the market purchases of these inputs are cut by farmers. Due to the possible violation of assumption 5.1, insecticides and fungicides will be dropped out of return computation. If the real impact on demand for insecticides and fungicides is not substantially different from our estimates, then the computed return on investment will be underestimated (as funds saved by cutting of market purchases of those factors will not be considered). Therefore the estimated economic return might be considered to be conservative.

[4]This is true for any size of package of inputs. The contain of basic package has been described in section 2.2.

5.3 Labor inputs

The second group of inputs for which the demand is expected to change in response to program participation are the labor inputs. The reason for this has been already mentioned in section 5.1 - increased requirements on workforce during input application and during the harvest as the output levels are higher. Further there might be a shift between different kinds of labor form based on the recommendations of CAA promoters.

In Ghana, like in the most of the developing countries, agriculture is very labor intensive. There are several different forms of labor; for purpose of this thesis divided into 5 groups - household labor (men, women or children living in the household), annual labor (on a daily basis, e.g. abunu or abusua forms[5]), contract labor - hired by day or by task and nnoboa labor (special labor sharing agreement). CAA recommends to use household and nnoboa labor (between CAA group members) as the optimal cost-minimizing strategy. Further advantage of nnoboa contracts is that the members are then more likely to share their experience about the application of hi-tech inputs and CAA program generally. Another advantage would be the improvement of monitoring between group members.

The statistics of the usage of individual forms of labor for non-members and members in all three examined seasons are displayed in tables 5.5 and 5.6 respectively. Only in the case of two forms of labor - annual and nnoboa - the usage of members dominates the usage of non-members. Another issue which is worth attention is a huge decline in usage of household and nnoboa labor for members in season 2008/2009. One of the possible explanations is that labor inputs are more likely to suffer from measurement error. Nevertheless, this decline corresponds to the decline in both output and yields as depicted in figures 4.1 and 4.2. While the decline in output could be possibly explained by a demand shock (e.g. due to global financial crisis), the decline in yield could not. In the case of yield the causality seems to go in the opposite direction - suboptimal use of labor led to the decline in yield. Examination of the true cause of this decline is not within the scope of this text.

To estimate the impact of the CAA program on the labor input demand, model 5.2 and the same estimation strategy as in sections 4.3 and 5.2 is used. There are, however, two small differences. First, it is possible to measure labor

[5]Abunu - "sharecropping agreement, tenant gives half of produce to owner", Abusua - "sharecropping agreement, tenant gives third of produce to owner" (Sarpong 2006)

Non-members								
	Household days		Annual days		Contract days		Nnoboa days	
Season	mean	median	mean	median	mean	median	mean	median
2007/2008	90.27	53	44.32	0	35	3.5	24.94	0
2008/2009	89.14	64.5	21.38	0	45.85	12	19.95	0
2009/2010	125.77	73.5	28.41	0	42.65	15	19.75	0

Table 5.5: Labor demand - in days

Members								
	Household days		Annual days		Contract days		Nnoboa days	
Season	mean	median	mean	median	mean	median	mean	median
2007/2008	115.57	74.5	36.89	0	49.33	13	33.27	0
2008/2009	86.36	55.5	21.70	0	56.94	8	18.78	0
2009/2010	133.99	84	17.57	0	33.99	15	32.27	12

Table 5.6: Labor demand

inputs in two alternative ways - either as the number of employed laborers in those forms of labor contract or as the number of work-days of the labor used. In this thesis (like in Zeitlin et al. (2009b) or Zeitlin et al. (2009a)) the number of laborers is used, because it is expected to be more precise (or better said, less imprecise).

Second, we enrich the set of control variables by including a new one - number of adults in a household. This serves as a proxy for the supply of household labor by the household. It seems reasonable to assume that more populous households are on average more likely to use household labor. Moreover it is expected to have an impact on the usage of other forms of labor input due to the substitution effect between individual forms.

The results of the estimation are displayed in table 5.7.

For the purpose of consistency in computation for different length of participation, only those labor inputs with sufficient significance in all three treatment groups (by length of treatment) will be included in ther computation of returns. The results of the estimation are very imprecise, which might be caused by measurement errors. The only type of labor close to conventional levels of significance is nnoboa labor (highest p-value equal to 0.111). Further the impact is relatively huge in magnitude hence the estimated impact on nnoboa labor demand will be included in the computation of economic returns.

Avg. Treatment Effect					95% Confidence Interval	
Household laborers	Coef.	Std. err.	t	p-value	lower	upper
1st year	.9688461	.9246945	1.05	0.298	-.8728426	2.810535
2nd year	1.989507	1.241538	1.60	0.113	-.4832303	4.462243
3rd year	-1.221133	1.809315	-0.67	0.502	-4.824697	2.382431
Annual laborers						
1st year	.1094015	.3751162	0.29	0.771	-.6377071	.8565101
2nd year	.7918318	.4314413	1.84	0.070	-.067458	1.651122
3rd year	.3993564	.560985	0.71	0.479	-.7179421	1.516655
Contract l., by day						
1st year	-.0928042	.8051762	-0.12	0.909	-1.696452	1.510843
2nd year	-.1373956	1.164706	-0.12	0.906	-2.457108	2.182317
3rd year	1.195162	1.528213	0.78	0.437	-1.848538	4.238862
Contract l., by task						
1st year	-1.10486	.4804266	-2.30	0.024	-2.061713	-.1480077
2nd year	-.7446447	.4506247	-1.65	0.103	-1.642142	.1528523
3rd year	-.5401985	.5490441	-0.98	0.328	-1.633715	.5533175
Nnoboa laborers						
1st year	4.104864	1.091745	3.76	0.000	1.930465	6.279262
2nd year	2.157824	1.33808	1.61	0.111	-.5071926	4.822841
3rd year	3.497484	1.534414	2.28	0.025	.4414344	6.553534

Table 5.7: Demand estimates - labor inputs

5.4 Economic returns computation

In this section, economic return on investment (CAA program) is computed using formula 5.1. Still, there are a few things essential for the evaluation of economic impacts, which have not been investigated yet. First, the average size of the CAA loans is not the same for 1,2 and 3 year members. Further, the package received by farmers in these subgroups is different. As a consequence, the impact on market purchases of fertilizer has to be adjusted for each group separately. Moreover, it is possible that not only the amount of laborers employed in response to the program participation varies across different groups, but also the average number of days they spend might differ. For a sake of simplicity, we use prices both of output and inputs from season 2007/2008 which were already used in Zeitlin et al. (2009b), Zeitlin et al. (2009a). This also allows easier comparison of economic returns in different treatment groups.
The prices were following: For 62.5 kg bag of cocoa farmers received on average 75 New Ghana Cedis (GHS). Fertilizer cost 15 GHS per bag. As in Zeitlin et al. (2009b), opportunity cost for own work is used to evaluate cost of labor under nnoboa agreement resulting in 2.5 GHS daily wage.[6]

The average size of the package obtained for groups of treated for 1,2 and 3 years was approximately 2.2, 2.7 and 3.5 acres worth respectively.[7] The average

[6]This is based on the fact that under nnoboa agreement farmer is obliged to provide the same amount of work as the other contracting side.

[7]The average size of the package obtained for farmers treated for one year exceeds 2

number of days nnoboa laborers have worked for individual treatment groups are approximately 4.75, 3.6 and 4.9 respectively. This leads to the following average gross profits (without the cost of investment):

$$\pi^G_{1y} = 472GHS \quad \pi^G_{2y} = 872GHS \quad \pi^G_{3y} = 846GHS$$

And average returns on investment are:

$$ROR_{1-year} = 55\% \quad ROR_{2-years} = 133\% \quad ROR_{3-years} = 74\%$$

The results support the hypothesis of increasing economic returns on program participation only partially. The fact that the economic returns (both in absolute and relative values) in the 2nd and the 3rd year of program participation exceed the returns in the 1st year supports our hypothesis. Nevertheless, the decline (both in absolute and mainly in relative values) contradicts the hypothesis. The main source of extraordinarily high returns in the 2nd year of treatment is the highest productive impact amongst groups; the changes in factor demand are small in magnitude therefore not likely to influence the outcome substantially. Possible explanation for such extreme agronomic returns were discussed in section 4.4. Most probable seems to be a possible measurement error in output (and other key control variables in the model).

acre package which should be provided to farmers in the first year. This can be caused by administration mistakes of CAA promoters, possible individual exceptions for some farmers or most likely due to the measurement errors. These errors might be partially resulting from necessary cleaning as the data about received dose for substantial part of the dataset were missing or corrupt. The cleaning procedure is described in appendix.

Chapter 6

Retention

6.1 Retention overview

In the previous chapter high positive returns on program participation have been estimated. Furthermore, the returns in later years are likely to be higher than in the first year of membership. Despite this fact, retention rates are lower than expected - namely 28% in season 2007/2008, 18% in season 2008/2009 and 14,4% in season 2009/2010. This is negative in two ways - those who leave are not participating in the CAA program for next season and even in the case they rejoin the program their treatment (nutrient stock accumulation) is interrupted. As a consequence it is the policy-maker´s interest to investigate this problem.

There are few possible explanations why should farmers leave the program. Suri (2006) and Zeitlin et al. (2010) pointed to the problem of heterogeneity in returns on adoption of agricultural technologies which are masked by average treatment effect. As a result, there might be a portion of farmers whose individual returns on program are close to zero or even negative. Nevertheless, farmers might believe before joining the program that their returns will be positive and after earning experience about their individual returns they might update their beliefs about their idiosyncratic returns and decide to leave. This was described as persistent heterogeneity. As an example of transient shock which can affect participation decisions has been mentioned depletion of buffer of savings highlighted in Dercon and Christiaensen (2007). With saving being depleted, farmers might not be willing to sustain their use of agricultural technology, which is considered to be risky.

The retention in CAA program has been already studied by Zeitlin et al. (2009b) and Zeitlin et al. (2010). They constructed a simple probit model in which they explained program retention by a dummy variable denoting farmers having problem with the repayment of their loan, dummy variable for farmers in whose group some of the other members have had repayment problem (in the latter case) and different proxies for individual return on program. Proxies used were change in $ln(cocoa)$ and two period change in $cocoa$.[1] Both individual's or peer's problem with complete repayment in time decreased the probability of staying in the program. On the contrary, they found positive relationship between individual return on program and probability of staying in the program.

6.2 Empirical evidence

In order to study retention of 2nd and 3rd year members as well, different set of controls for individual return has to be introduced. Simple 1 (2) period change in output would not make much sense. For 3rd year member such measure would represent comparison between output in 1st (2nd) year and between output in 3rd. This hardly captures farmers beliefs in return on program participation. The problem is, however, that the size of the information set available to 1st, 2nd and 3rd year members differs. While farmers being treated for 1 year have only one realization of experiment regarding their individual return, farmers being treated for more years have more realizations. Unfortunately, the size of the subset of 2nd and 3rd year members is small and their retention will be fully explained by their and their peers' repayment problems. Due to their possible correlation with measures of individual returns, models using only such sub-samples will not be consistent (if those variables are omitted). As a solution we introduce one measure for all treatment groups which account for different set of experience of idiosyncratic returns of farmers to program participation. We construct it in the following way:

$$\begin{aligned} IR = & 1\{\text{1-year member}\}(y_{i,t} - y_{i,t-1}) + \\ & 1\{\text{2-year member}\} \sum_{j=0}^{1} w_{2j}(y_{i,t-j} - y_{i,t-2}) + \\ & 1\{\text{3-year member}\} \sum_{k=0}^{2} w_{3k}(y_{i,t-j} - y_{i,t-3}) \end{aligned} \tag{6.1}$$

[1] This was based on the high correlation of two period change in output and estimated quantile treatment effect.

This can be basically describes as weighted sum of differences of output within program and pre-program output. The weights are chosen arbitrary (with $w_{20} + w_{21} = 1$ and $w_{30} + w_{31} + w_{32} = 1$), though we assume that farmers current decision about future program participation is more affected by newer experience hence $w_{20} \geq w_{21}$ and $w_{30} \geq w_{31} \geq w_{32}$. As in other literature we include dummy variables for repayment problems of both farmers themselves and their peers. Finally we include in the model controls for specific season and observable farmer characteristics (gender, education and age). We estimate the following model using probit method:

$$P(Fmember = 1) = \Phi(\alpha RP + \beta PRP + \gamma IR + \boldsymbol{\delta X}) \quad (6.2)$$

Where RP denotes repayment problems of a farmer, PRP denotes repayment problems of a farmer's peers, IR is individual return defined by equation 6.1 and $\boldsymbol{X}$ is a vector of other control variables.

Testing the model 6.2 using three different weighting schemes [2]. The results of estimation under all the weighting schemes are reported with clustered robust standard errors in table 6.1 (reporting only variables of interest).

					95% Confidence Interval	
Equal weights	Coef.	Std. err.	z	p-value	lower	upper
Repayment problem	-1.17771	.4806275	-2.45	0.014	-2.119723	-.2356974
Peer repayment problem	-1.647167	.4357513	-3.78	0.000	-2.501223	-.7931098
Individual return	.0002316	.0001007	2.30	0.022	.0000341	.000429
Only current						
Repayment problem	-1.190272	.4766984	-2.50	0.013	-2.124584	-.2559601
Peer repayment problem	-1.612858	.4319368	-3.73	0.000	-2.459439	-.7662775
Individual return	.0001373	.0000784	1.75	0.080	-.0000164	.0002909
Decreasing weights						
Repayment problem	-1.1824	.4788162	-2.47	0.014	-2.120862	-.2439372
Peer repayment problem	-1.628376	.4332382	-3.76	0.000	-2.477507	-.7792447
Individual return	.0001989	.0000953	2.09	0.037	.0000122	.0003855

Table 6.1: Program retention estimates - probit model

The results are in accord with results found by Zeitlin et al. (2009b), Zeitlin et al. (2010) - negative effect of both own and peers' repayment problems

[2]

1. Only most recent experience (i.e. $w_{20} = w_{30} = 1$ and $w_{21} = w_{31} = w_{32} = 0$)
2. Equal importance of each year experience (i.e. $w_{20} = w_{21} = \frac{1}{2}$ and $w_{30} = w_{31} = w_{32} = \frac{1}{3}$)
3. Decreasing impact for older realizations: $w_{20} = \frac{2}{3}$, $w_{21} = \frac{1}{3}$, $w_{30} = \frac{4}{7}$, $w_{31} = \frac{2}{7}$ and $w_{32} = \frac{1}{7}$.

and positive effect of individual return on program. Estimates are significant under all three different weighting schemes. Positive effect of individual return supports the heterogeneity in returns amongst farmers. This issue is not within scope of this thesis.

Chapter 7

Conclusion

The question of returns on adoption of agricultural technology is naturally of a big interest in development economics and has been studied a lot. Despite the high average returns on adoption of agricultural technology, the rate of adoption and sustained usage of such technologies was relatively low. As a consequence, many studies tried to find possible explanation for this phenomenon. Examples of possible explanation provided in the literature are: Heterogeneity in returns across farmers (Zeitlin et al. 2010), (Suri 2006), risk aversion (Dercon and Christiaensen 2007), behavioral causes like procrastination (Duflo et al. 2009) and other.

This thesis examines the CAA program providing hi-tech agricultural inputs on credit with a joint liability to groups of cocoa farmers in Ghana. It contributes by presenting and estimating a model for the identification of program impacts in the 2nd a the 3rd year of sustained membership (besides 1st year). Further it adjusts the model of CAA program retention provided in literature (Zeitlin et al. 2009b), (Zeitlin et al. 2010). The reasons why there should be higher agronomic and economic return on CAA program are the following: First, nutrient stock should accumulate during longer treatment and therefore result in higher output. Second, the alleviation of credit constraint for farmers should have higher effect during longer membership because there is a possibility of extension of the size of the loan provided by CAA program.

Empirical evidence provided in this thesis has led to the following conclusions: Productive impacts (in absolute values) are substantially higher in later years of program participation. On the contrary, substantial difference in agronomic returns on loan has not been found. Unless there is a very significant cut in market purchases of inputs in later years, the only channel through which

the program increases production in later years of program participation is by extension of the size of the loan.
Gross profit resulting from CAA program participation is higher in later years than in the 1st. Based on the evidence regarding agronomic returns this can be mainly attributed to the extension of the size of the loans. Unlike the productive return, the economic return in later years of CAA program participation is higher than during the first year. This might be attributed to more substantial decline in market purchases of agricultural inputs in later years of program participation.
The evidence confirms the conclusion about program retention determinants found in the literature (Zeitlin et al. 2009b), (Zeitlin et al. 2010) - the negative impact of individual's or group peer's repayment problem on probability of program retention and positive relation between farmer's idiosyncratic return on program and probability of staying in it.

From a policy perspective the results mentioned above have a few important implications. Relatively high drop-out rates present a serious problem, because they prevent farmers from enjoying larger economic returns in the later years of program participation. Further it is possible, that some farmers, whose realized economic returns in the 1st year of program participation were negative (or close to zero), will reach positive returns in later years. As a consequence dealing with high drop-out rate is of a vital interest of a policy maker.

Some of the possible measures which would lead to lower drop-out rate and/or more common adoption of larger package sizes were mentioned in Zeitlin et al. (2009b). First was the provision of weather and other micro-insurance. This will help farmers overcome the transient shocks to their return on program and therefore lower the drop-out rate. Another measure mentioned was allowing for extension of the size of the loan for individual farmer above the size provided to group (on individual liability). Further it would be useful to inform farmers about higher returns in later years on program participation.

Bibliography

Conley, T. G. and Udry, C. R. (2005). Learning about a new technology: Pineapple in Ghana. *Proceedings*.

Dercon, S. and Christiaensen, L. (2007). Consumption risk, technology adoption, and poverty traps : Evidence from Ethiopia. Policy Research Working Paper Series 4257, The World Bank.

Duflo, E., Kremer, M., and Robinson, J. (2008). How high are rates of return to fertilizer? Evidence from field experiments in Kenya. *The American Economic Review*, 98(2).

Duflo, E., Kremer, M., and Robinson, M. (2009). Nudging farmers to use fertilizer: Theory and experimental evidence from Kenya. Working Paper 15131, National Bureau of Economic Research.

Graybeal, D. K. (2000). Variation in marginal response to nitrogen fertilizer between locations. *Journal of Agricultural and Applied Economics*, 32(02).

Hartemink, A. E. (2005). Nutrient stocks, nutrient cycling, and soil changes in cocoa ecosystems: A review. *Advances*, 86(1):227–253.

ICCO (2010). The world cocoa economy: Past and present. Technical report, ICCO.

Kolavalli, S. and Vigneri, M. (2011). Cocoa in Ghana: Shaping the success of an economy. *World Bank*, pages 1–26.

Matsumoto, T. and Yamano, T. (2010). The impacts of fertilizer credit on crop production and income in Ethiopia. GRIPS Discussion Papers 10-23, National Graduate Institute for Policy Studies.

Quartey, E. (2007). *COCOBOD presentation*.

Ricker-Gilbert, J., Jayne, T. S., and Black, J. R. (2009). Does subsidizing fertilizer increase yields? Evidence from Malawi. 2009 Annual Meeting, July 26-28, 2009, Milwaukee, Wisconsin 49532, Agricultural and Applied Economics Association.

Sarpong, G. A. (2006). Improving tenure security for the rural poor Ghana – country case study. Technical report, FAO.

Suri, T. (2006). Selection and comparative advantage in technology adoption. Working Papers 944, Economic Growth Center, Yale University.

Zeitlin, A. (2010). Producer networks and technology adoption in Ghana. *CSAE Working Paper*.

Zeitlin, A., Caria, S., and Teal, F. (2009a). Impacts of group-based microfinance in agriculture : Evidence from Ghana. *CSAE Working Paper*.

Zeitlin, A., Caria, S., and Teal, F. (2009b). Improving productivity through group lending: Report on the impact evaluation of the Cocoa Abrabopa initiative. *CSAE Working Paper*.

Zeitlin, A., Caria, S., Teal, F., Dzene, R., Jansky, P., and Opoku, E. (2010). Heterogeneous returns and the persistence of agricultural technology adoption. *CSAE Working Paper*.

Zerfu, D. and Larson, D. F. (2010). Incomplete markets and fertilizer use: Evidence from Ethiopia. *Africa*, (5235):19 pp.

Appendix A

Appendix

A.1 Tables and figures

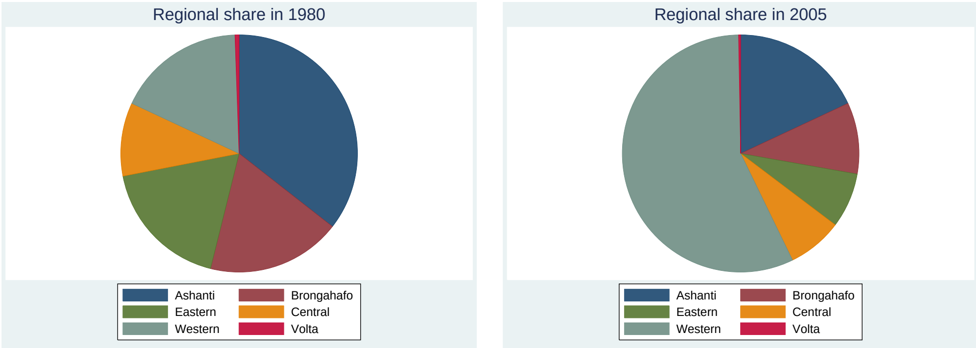
Figure A.1: Map of Ghanaian regions

Source: http://www.world-geographics.com/maps/africa/map-of-regions-in-ghana/

	Non-members		Members	
Season	mean	median	mean	median
2007/2008	49.27	50	49.19	48
2008/2009	49.56	49	50.49	52
2009/2010	52.06	52	49.63	49

Table A.1: Age distribution

Figure A.2: Regional share in cocoa production



Year	Output (tonnes)	Yield (kg/ha)
2001	389,772	2,886
2002	340,563	2,850
2003	496,846	3,313
2004	736,976	3,685
2005	599,318	4,000
2006	740,458	4,000
2007	614,532	4,200
2008	680,781	3,735
2009	710,642	4,291
2010	650,941	3,889

Table A.2: Total production and yield evolution

Cocoa kg	Coef.	Robust Std. err.	t	p-value	95% Confidence Interval lower	upper
nFmember1y	1043.291	332.9894	3.13	0.002	380.085	1706.497
nFmember2y	65.81477	288.0604	0.23	0.820	-507.9073	639.5369
nFmember3y	630.5587	309.4437	2.04	0.045	14.2481	1246.869
Fmember1y	435.2629	169.04	2.57	0.012	98.59066	771.9352
Fmember2y	762.6809	234.7061	3.25	0.002	295.2231	1230.139
Fmember3y	1140.111	292.5941	3.90	0.000	557.3599	1722.863
Fmember4y	1167.394	252.9217	4.62	0.000	663.6566	1671.131
Early ad.	-502.779	213.7182	-2.35	0.021	-928.4356	-77.1224
2nd year ad.	-122.8323	199.9079	-0.61	0.541	-520.9834	275.3188
3rd year ad.	-529.0187	280.5783	-1.89	0.063	-1087.839	29.80147
Late ad.	389.1283	371.2523	1.05	0.298	-350.2848	1128.541
Former mem.	867.478	233.93	3.71	0.000	401.566	1333.39
Re-joiner	273.3021	248.6628	1.10	0.275	-221.9527	768.5569
Education	171.8349	85.19918	2.02	0.047	2.145995	341.5237
Female	-247.5803	90.0705	-2.75	0.007	-426.9712	-68.18931
Hybrid land	152.4614	25.86736	5.89	0.000	100.9421	203.9807
Non-h. land	165.8904	22.06153	7.52	0.000	121.951	209.8297
Constant	534.3008	137.2181	3.89	0.000	261.0074	807.5943

R-squared: 0.3180
observations: 1017

Table A.3: Output estimation results

Avg. treatment effect Cocoa kg	Coef.	Std. err.	t	p-value	95% Confidence Interval lower	upper
1st year	444.8604	161.0553	2.76	0.007	124.091	765.6297
2nd year	776.8416	273.9762	2.84	0.006	231.1707	1322.512
3rd year	659.031	201.5104	3.27	0.002	257.6883	1060.374

Table A.4: Output ATE, only early adopters, one outlier removed

Avg. treatment effect Cocoa kg	Coef.	Std. err.	t	p-value	95% Confidence Interval lower	upper
1st year	342.4267	148.2652	2.31	0.024	47.1309	637.7225
2nd year	579.0384	224.0026	2.58	0.012	132.8987	1025.178
3rd year	588.3416	190.9251	3.08	0.003	208.0812	968.602

Table A.5: Output ATE, yield $\in (30; 2000)$

A.2 Cleaning procedure - data on received dose

Data regarding the dose received were missing or apparently wrong on a large scale. In order to be able to compute average dose received, the data had to be cleaned. Cleaning was based on the following considerations:

- Farmers answer about the size of the loan in other than current season is correct. Based on this, values were filled in according to other farmer's answers in other years. In the case of conflict of two or three farmers answers the maximum dosage was used. If this correction will be incorrect for some farmers, this will only result in underestimation of economic impacts of the CAA program.

- Farmers are not expected to deviate from common size of the loan within group. Any exceptions or administrative errors are considered to be rare. Based on this, the missing values were filled in as a mode of dose received by group members.

- Farmers outside the program should not receive any package. This, however, does not affect any computations.

These steps have been done in the given order.