Master thesis for the Master of Economics degree

Debt and Poppy Cultivation

Driving factors behind Afghan opium production

Fredrik H. Willumsen

November 13, 2006

Department of Economics

University of Oslo

Preface

Supervisor: Kalle Moene, University of Oslo.

Summary

In this master thesis I intend to study opium production in Afghanistan, and identify important drivers behind the opium production. The main aim is to test whether or not Afghan opium production is debt–induced.

A claim often found in the literature on Afghan opium production is that the production of opium is debt–induced. In the first part of the thesis I provide a theoretical rationale for why this is the case, using a dynamic model of the cropping choice of a utility maximizing household. The main finding is that optimal effort devoted to opium production is increasing in the level of debt.

In the second part of the thesis I describe the data set, and put up correlation tables for opium production and debt. The findings are consistent with the literature—the conditional probability of producing opium given debt is significantly higher than the conditional probability of producing opium given non–negative wealth.

But correlation does not necessarily imply causation. In the third part of the thesis I therefore estimate reduced form versions of the theoretical model, calculating the choice probabilities for different cropping strategies. Here the findings diverge from the literature—when controlling for price incentives, eradication risk, social class, and regional–specific fixed effects, the conditional probability of producing opium is independent of debt.

To explain this finding I, in the final part the thesis, show that heterogeneity in moral costs may create two subpopulations of opium farmers. The first, the "opportunists", produce opium for the unrivalled profit, while the second, the "moralists", produce opium out of necessity. Utilizing the theoretical model with moral costs, I find that the "moralists" either will produce opium at maximum capacity or not at all, while the "opportunists" will have a production level somewhere in between. Taking this into account when calculating choice probabilities, our model fits the data better: for the "moralists" debt is an important determinant of opium production, while for the "opportunists" debt is unimportant.

Acknowledgements

This master thesis is written as a part of a bigger project on the underdevelopment of Afghanistan, led by Kalle Moene. I will like to thank my supervisor Kalle Moene for involving me in this project, which is a part of his engagement at the Centre for the Study of Civil War at the Peace Research Institute of Oslo (PRIO).

Contents

Preface i

Contents ii

1 Introduction 1

2 The Afghan opium economy 4

2.1 The salaam system 6

2.2 Farmers' incentives to produce opium 7

3 A model of debt–induced opium production 9

3.1 Two–period model without uncertainty 9

3.2 Infinite horizon model with uncertainty 14

3.3 A model with moral costs 17

4 Description of the data 21

5 Empirical strategy and results 26

5.1 Conditional probability of producing opium as a function of debt 28

5.2 Tobit regression 33

5.3 Discriminating between "moralists" and "opportunists" 33

5.3.1 Test of the implied correlation structure for the error terms 38

6 Conclusions 42

References 44

A Computer code 47

A.1 Binary logit 47

A.2 Tobit regression 47

A.3 Multinomial logit 48

A.3.1 Nested multinomial logit 48

1 Introduction

The purpose of this master thesis is to describe the Afghan opium farmers and examine the incentive structure the farmers face: what characterize the farmers that choose to produce opium and what are the main factors that motive the farmers to produce opium? We are especially interested in the effect of debt on the choice of cropping strategy.

Why is this an important topic? There is a widespread belief that anti–drug efforts have to be aligned with the incentives of the farmers to be successful. Hence, if debt is an important determinant of opium production, the Afghan government should focus more on alternative development than eradication efforts. But despite of the enormous focus on the link between debt and opium production in the literature there has, to my knowledge, not yet been done a thorough statistical analysis of the question. Studies in this line of research are often of a more qualitative nature, more focused on in–depth interviews than quantitative analysis.

The questions we address are the following:

- Is opium production debt–induced? And for what parts of the population is opium production debt–induced?
- To what income group(s) do the farmers that produce opium belong?
- What is the relative importance of price incentives?
- What is the effect of the government of Afghanistan's eradication program?

The data set used is from a survey conducted by the United Nations, Office on Drugs and Crime (UNODC). The data set is thoroughly described in UNODC's report from the survey (UNODC, 2004). UNODC (2004) also examine some of the questions we are interested in, however they mostly analyze and report simple correlation schemes. We are more interested in doing a more thorough statistical analysis of the cropping decisions of the Afghan farmers, going a step beyond reporting correlations. We use discrete choice methods to analyze what factors are important determinants for opium production—taking into account that explanatory variables may be correlated, which can make inference from simple correlations biased.

The thesis is organized as follows. The first part presents a simple two–period model for the cropping decision, where the farmer/household acts as a utility maximizer subject to resource constraints. The model shows that opium production will be an increasing function of the level of debt and, provided that the debt is large enough, initial debt will affect the cropping profile of the farmer in several consecutive periods. In addition, because of the salaam debt system, production will be shifted towards production today relative to production next period if the farmer has debt, while not if the farmer has positive wealth. To avoid the effect of the induced stop at time 2, and since opium production obviously is subject to stochastic disturbances, we expand our model to an infinite horizon model with uncertainty. The finding here is equal to the finding in the two–period model, effort devoted to opium production is increasing in the level of debt. Moral costs related to the production of opium is also an important determinant of opium production. When this is introduced into the model we find that farmers with a high moral cost term will have more discontinuous shifts in the optimal strategy than other farmers. The farmers with a high moral cost term will, when their level of debt passes some individual threshold determined by their moral cost term, supply very high effort to be able to repay their debts. The effort will be increasing in the moral cost term.

In the second part of the thesis I describe the data set, and put up correlation tables for opium production and debt. The findings here are consistent with the literature—the conditional probability of producing opium given debt is significantly higher than the conditional probability of producing opium given non–negative wealth.

In the third part of the thesis I estimate reduced form versions of the theoretical model, calculating choice probabilities for different cropping strategies. Here the findings diverge from the literature—when controlling for price incentives, eradication risk, social class, and regional–specific fixed effects, we find no significant effect of debt on the conditional probability of being an opium producer.

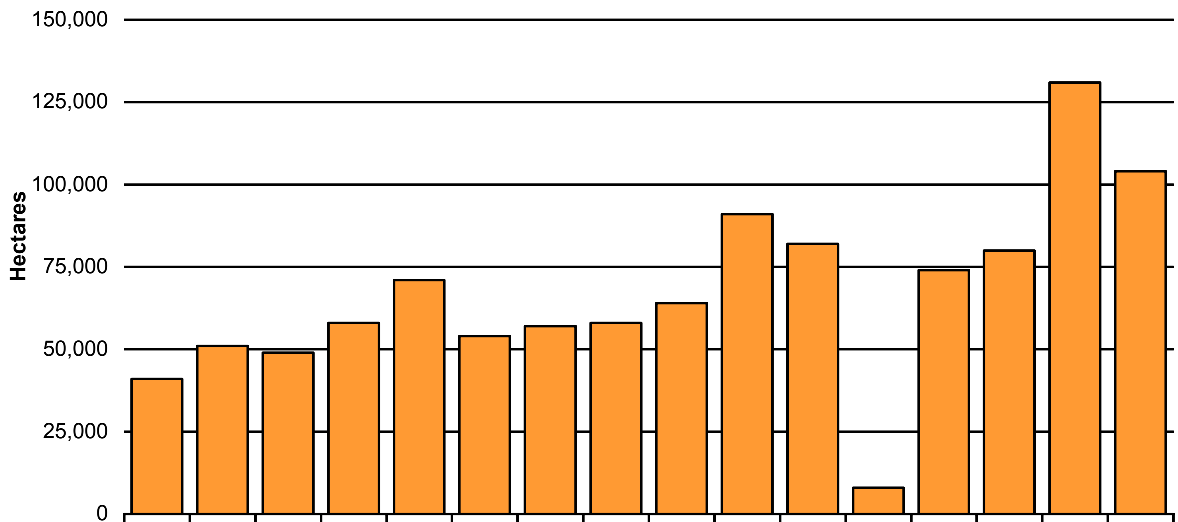
To explain these findings I, in the final part the thesis, show that heterogeneity in moral costs may create two subpopulations of opium farmers. The first, the "opportunists", maximize profit, while the second, the "moralists", produce opium out of necessity. We also find that the two groups will have a different cropping strategy, which makes us able to distinguish them empirically: the "moralists" will produce zero opium until their accumulated debt reaches a tipping point, and when this point is reached they will produce at maximum capacity. The "opportunists" will find it optimal to diversify their production. Taking this into account when calculating choice probabilities, our model fits the data better: debt is an important determinant of opium production for the "moralists", while unimportant for the "opportunists".

This thesis is a part of a research project on Afghanistan, led by Kalle Moene, where we look at opium production as a special case of a resource curse. It is a *curse* in the sense that production of opium works as a poverty trap—poverty is the reason why the farmers produce opium, and at the same time opium production is a hinder for alternative development and industrialization of the rural parts of Afghanistan, hence it is also the reason why the farmers are poor. It is a *resource* curse since the Afghan soil is extremely well suited for production of opium. We hypothesize that the opium traders induce opium production by lending money to poor farmers at salaam terms, terms which imply that the farmers must repay the loan in kind. To be able to obtain the loan, or at least to be able to repay it, the farmers will have to produce opium. In this way the traders have an instrument as moneylenders to control what the farmers produce, without producing opium it is much harder for farmers to obtain loans. A strong implication of this general hypothesis is that farmers that have debt will have a different cropping pattern than other farmers, and the investigation of whether this is the case or not is a central part of this thesis.

The thesis draws on a (small) strand of literature focusing on the factors that drive opium production in Afghanistan. Several authors have attributed the increase in opium production after 2001 to the persistent high prices of raw opium (see for example Molla (2003) and UNODC (2003a)), while others point out a more diversified set of incentives for farmers to increase their production (Mansfield, 2001, 2004b, 2005; Mansfield and Pain, 2006; UNODC, 2004). Clemens (2005) estimates a structural model to test the effect of different policy regimes on opium production. This master thesis draws especially from Mansfield (2004b, 2005) and UNODC (1999b), while doing more advanced statistical analysis like Clemens (2005). However, as we have micro–level data, we put weight on the moral costs of production and the effect of debt. Clemens (2005) uses macro–level data on opium production, focusing primarily on price incentives, labor constraints, stigma of opium production and eradication risk.

2 The Afghan opium economy

Afghanistan is *the* major opium producing country in the world. Newly released figures for production in 2005, reported in the World Drug Report 2006 (UNODC, 2006), show that Afghanistan produces 89 % of the world's total produced quantity of opium, making the country the sole supplier of opium to the world market. However, the World Drug Report 2006 also reports that opium production in Afghanistan in 2005 has decreased by nearly 20 % relative to 2004, a significant decrease which comes after three consecutive years of growth in both produced quantity and hectares of land devoted to opium production. As we can infer from this, and from Figure 1, the amount of opium produced and the amount of land devoted to opium production are not static sizes, they exhibit substantial variation both over time and over space. This implies that farmers are going in and out of production of opium, a variation which we also find in our data set. This variation is important for our analysis, as it gives us the opportunity to pin down the values of the parameters in our model.



	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
Cultivation	41,000	51,000	49,000	58,000	71,000	54,000	57,000	58,000	64,000	91,000	82,000	8,000	74,000	80,000	131,000	104,000

Figure 1: Afghanistan opium poppy cultivation in thousands of hectares. Source: World Drug Report 2006

One question naturally arises: how did Afghanistan end up as the world's number one narcotic state? While there of course are several reasons for this, a quick look at recent historical events might prove valuable.

UNODC (2003b, page 10) identifies three major causes as to why the opium economy in Afghanistan developed:

- lack of effective government administration

- degradation of agricultural and economic infrastructure
- a war economy and related black marketeering

Afghanistan has been a war site ever since the Soviet invasion in 1979. Due first to the invasion, and later to the political vacuum that was created after the fall of the Soviet empire and the withdrawal of the Soviet forces, there has never been a truly effective government administration in Afghanistan. Not even today, five years after the US–led invasion in 2001, and with the presence and support of huge numbers of foreign military troops, the government can be said to have control with all areas of the country. This lack of effective government control has given drug traffickers and other criminals the opportunity to develop an illegal economy, consisting of trade with arms, drugs and financing of terrorist activities (UNODC, 2003b).

A consequence of the wars and conflicts is that most of the important agricultural and economic infrastructure has been destroyed. The destruction of the infrastructure has made opium production even more attractive, since opium production does not face the same constraints as other agricultural production. Opium is for example storable, hence the lack of transportation possibilities and roads is less severe than if the farmer produced vegetables that had to be sold at a market right after harvesting. Opium is also less susceptible to lack or irregularities in the supply of water (Mansfield, 1999), a comparative advantage when most of the irrigation systems are destroyed.

Another consequence of the conflicts is that the financial sector has ceased to exist, and an informal financial sector has developed (UNODC, 2003b). Two major problems have emerged from this. First, the number of loans taken up for investments has dropped, two thirds of the loans taken up are now used for social needs, such as food and clothing (ACBAR, 2000). Second, the preferred collateral is opium, which gives extreme incentives to produce opium (Mansfield, 2001). Opium becomes a poverty trap—farmers produce opium to be able to obtain credit to survive, but since they are producing opium, and hence are contributing to the informal economy, investments in economic infrastructure and the formal sector are crowded out.

Yet another important factor for the emergence of the opium economy was the Taliban, which can be said have emerged as a response to the lack of a state apparatus in the beginning of the 1990's. This movement in fact encouraged opium production, at least from 1997 and onwards (Rashid, 2002), as they realized that opium production could fund their own activities. And even though they strictly enforced laws against other forms of criminal activities (the use of opium was for example prohibited, and the debarment was extremely efficiently enforced) the production of opium for use outside Afghanistan was not considered to be anti–Islamic.

Also several other reasons as to why the narcotic state emerged can be provided. The crackdown on production of opium in Iran and Pakistan during the 1970's and the 1980's was important, as it pushed illegal activities into Afghanistan, where the lack of a state apparatus gave wide opportunities for opium traders and other criminals. Afghanistan place on the map is also important; situated in Central Asia, it is possible to smuggle opium to markets in Asia, Russia, and Europe. Finally, the Afghan soil is extremely well suited for opium production. The average yield of opium per hectare land in Afghanistan is 45 kg, compared to 7 kg/ha in Laos and 11 kg/ha in Myanmar (UNODC, 2003b).

As a tentative conclusion, it is reasonable to believe that destruction due to war and conflict along with the lack of a controlling central government, are major explanatory factors behind the growth of the Afghan opium economy. In the war economy, opium has become an integral part of the livelihood strategies of many of the rural communities in Afghanistan (UNODC, 2003b).

2.1 The salaam system

The salaam system is an informal credit system where the farmer obtains advances on a fixed amount of future agricultural production (UNODC, 1999b). The amount that the farmer receives is often half of the prevailing market price at the time when the farmer enters into the salaam arrangement. UNODC (1999b), for example, reports that, in a sample of 108 farmers from Qandahar and Shinwar, the price received by the farmers was on average 42 % of the value of the produced opium at the harvest time. This factor, labeled α below, is found to vary over time and space, and it also exhibits a seasonal pattern, being smaller when close to the harvesting time (UNODC, 1999b).

While this may seem to be an exploitive credit arrangement favoring the lenders on behalf of the farmers, the reason that this system evolved is of a quite different nature. The salaam system arose as a response to the Islam prohibition of interest on loans, and became popular in the Taliban period due to the Taliban's very restrictive interpretation of Islamic laws. The salaam system was introduced as a mean of providing credit, while not breaking Islam laws. Of course, since the salaam arrangement gives the farmers only half the value of the produce relative to harvest time there is a huge implied interest rate present. However, there is also a risk sharing between lender and borrower, because they both risk losing due to future price and harvest variation, and this risk sharing makes the salaam system legal under Islam; the lender does not take interest per se, he receives a compensation for taking risk.

Importantly, this is the ideal version of the salaam system. In Afghanistan today, where credit is needed to get through the winter, defaulting on a loan is not an option. Defaulting on a loan will imply that the farmer does not have the possibility of obtaining credit the next winter season, and hence the farmer will have difficulties surviving. If the farmer experience a crop failure, there exist two ways of repaying the loan. The first is to buy opium at the local bazaar, but if one farmer has experienced a crop failure there is a high probability that also other farmers in the area have experienced the same, and hence the supply of opium at the local bazaar may be non–existent, or opium may be available at a very high price. The other possibility is to reschedule the debt. The debt will then be rescheduled at salaam terms, which imply that the farmer will have to produce twice the amount of opium the next season to be able to repay the loan (These two alternative repayment strategies are documented in Mansfield (2001)).

As we can infer from this, the salaam system may lock farmers to opium production multiple periods after the initial loan was granted. In addition, as opium is considered to be the preferred collateral (Mansfield, 2001), we a priori should expect to find that debt is a strong determinant of opium production.

2.2 Farmers' incentives to produce opium

In a series of papers, David Mansfield (2004b, 2005, 2006) has identified important driving factors behind Afghan opium production. Mansfield (2004b) shows that price incentives alone are not able to explain the diversity in the cropping pattern in the 2003/2004 season, instead he focuses on food security and access to land and credit as important determinants of opium production. Mansfield also found that there was a growing confidence in the supply of wheat, and this made the farmers substitute production of wheat for own consumption for opium production. Further, accumulated debts and the opium–dependent salaam credit system are considered important determinants. Finally Mansfield finds that the eradication campaign initiated by the Afghan government has been counterproductive, the farmers who have experienced eradication increased the amount of land dedicated to opium production by more than the farmers who did not experience crop eradication. The rationale is that when the crops are eradicated, the farmer has to reschedule his loan, and due to the salaam credit system this will imply that he will have to produce even more the next season. There is however no discussion of whether these differences are statistically significant.

Mansfield (2005) finds that there in the 2004/05 season is likely to be a downturn in the number of households cultivating opium, and he attributes this to falling farmgate prices, concerns over food security, and an increased awareness among farmers of the risk of eradication. However, he still finds that households that experienced eradication last year, devote more land to opium production than other farmers.

Mansfield (2006) finds that eradication often targets the poorest, while those that have links to the authorities or have the financial strength to bribe escape the eradication measures. Again this will drive up the accumulated debts of the poor, contributing to increases in opium production.

UNODC (2004) reports that the two main reasons for farmers *not* to engage in opium production are that opium is against Islam and that opium production is illegal. Clemens

(2005), in a paper where he estimates the effects of different drug control policies using a partial equilibrium framework, finds that the stigma of producing opium is important to be able to account for the observed prices.

We will draw heavily on these identified drivers in the empirical part of this thesis.

3 A model of debt–induced opium production

In this section we show why opium production can be partly debt–induced, and that optimal opium production will be an increasing function of the level of debt.

We first model the cropping decision in a simple two–period model without uncertainty. We then introduce stochastics into the model, while we at the same time let the time frame go to infinity. Finally we consider the effect of moral costs on opium production.

The presence of the non–linearity in the budget constraint (equation (2)), the kink around zero due to the salaam debt system, makes the problem somewhat less straightforward to solve. The difficulty that arises, is that there are different cases to consider depending on the initial wealth W_0. The analysis that follows will therefore be somewhat tedious, as a thorough analysis requires that the optimal strategies for all the cases are identified. The intuition is however very simple. The higher the level of debt, the more utility the household get from producing opium. And given the salaam debt system, farmers will find it optimal to shift production to the present period relative to the next if they have debt, while not if they have positive wealth. For the model with moral costs, the important intuition is that farmers with a high moral cost term will have more discontinuous shifts in the optimal opium cropping strategy.

3.1 Two–period model without uncertainty

Let $u(c_t)$ denote a standard utility function, additatively separable over time, where c_t denotes level of consumption at time t, and the standard conditions $u'(c_t) > 0$ and $u''(c_t) < 0$ holds for all c_t. Let $g(x_t)$ denote the cost of producing x_t worth of opium, where $g'(x_t) > 0$ and $g'' > 0$. The household's objective is

$$\max_{\{c_t, x_t\}} \sum_{t=1}^{2} \beta^t \left[u(c_t) - g(x_t)\right] \tag{1}$$

$$\text{s.t.}$$

$$W_t = \begin{cases} x_t + W_{t-1} - c_t & \text{if } x_t + W_{t-1} - c_t \geq 0 \\ (x_t + W_{t-1} - c_t)\alpha & \text{otherwise} \end{cases} \tag{2}$$

$$c_t \geq \bar{c} \,\forall\, t \tag{3}$$

$$W_2 \geq 0 \tag{4}$$

$$W_0 \text{ given} \tag{5}$$

The constraints need some explanation. Constraint (2) is the intertemporal budget constraint. W_{t-1} denotes wealth inherited from $t-1$ (possibly negative), x_t the value of the farmer's production of opium and c_t consumption, all at time t. $\alpha > 1$ is the salaam parameter. Any debt taken up or rescheduled in period t, denoted b, will be transferred into a promise to produce αb worth of opium in period $t+1$. Anecdotal evidence suggests that $\alpha \approx 2$ (Mansfield, 2001). We here assume that the individual farmer is not in a position to lend money to other farmers at salaam terms (the α–parameter is only present when $x_t + W_{t-1} - c_t < 0$). Is this a reasonable assumption? We claim so: since the warlords and commanders are governing the opium industry it is highly unlikely that a farmer can become a creditor, lending to other farmers at salaam terms. The warlords and commanders want control over the production of opium, and they get this control through controlling the credit system, hence they are unlikely to be willing to introduce unnecessary competition in the credit market, deteriorating their own position as the sole provider of credit. Constraint (3) states that consumption must be above some exogenously given minimum level. Constraint (4) states that all debt must be repaid when period 2 is over, and since the utility function is globally increasing, constraint (4) will hold with equality. Constraint (5) says that the initial level of wealth is exogenously given. We do not model why some farmers are initially rich, while others are initially poor.

Let S denote the sequence of feasible choices $((c_1, x_1), (c_2, x_2))$, i.e. the set of sequences that fulfills all the constraints. We assume that S is not empty, i.e. we assume that the level of debt at $t = 1$ is sustainable. Let us further, for simplicity, assume that $\beta = 1$ and normalize $\bar{c} = 0$. We solve the problem backwards, starting at $t = 2$. Let $J(t, W_{t-1})$ denote the optimal value function at time t.

$$J(2, W_1) = \max_{c_2, x_2} \{u(c_2) - g(x_2)\} \text{ s.t. } \begin{cases} x_2 + W_1 - c_2 = 0 \\ c_2 \geq 0 \end{cases}$$

We solve for c_2 in the constraint, yielding the unconstrained optimization problem

$$J(2, W_1) = \max_{x_2} \{u(x_2 + W_1) - g(x_2)\}$$

First order condition for interior maximum, the intratemporal optimality condition, is

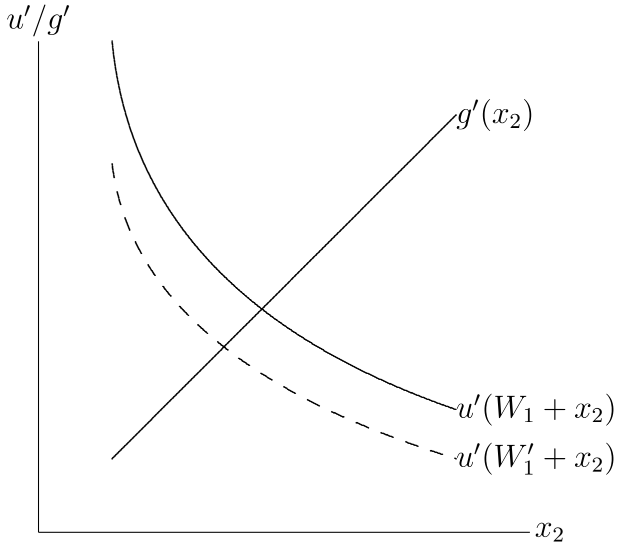
$$u'(x_2 + W_1) = g'(x_2) \tag{6}$$

Assuming an interior optimum exists, x_2 can be shown to be a decreasing function in W_1 provided that u is strictly concave and g is convex. More formally

Proposition 1 *Assume that $u'(c_t)|_{x_t=0} > g'(0)$. Then there exist a function $h : W_1 \to x_2$ that, provided that g is increasing and convex, u is increasing and strictly concave and an interior solution exists in the relevant domain of W_1 and the relevant range of x_2, uniquely determines the level of opium production. The level of opium production will be strictly decreasing in wealth.*

Proof. We prove the statement in Figure 2. A positive shift in W_1 shifts the marginal

Figure 2: Proof of Proposition 1 ($W_1' > W_1$)



utility curve inwards. The marginal cost curve is not affected by this shift. As long as the marginal utility curve is *strictly* concave, the curve will be downward sloping in x_2. And as long as marginal cost curve is convex (note that strict convexity is not needed), the marginal cost curve will be horizontal or increasing in x_2. It then follows that there always will be an intersection between the two curves, and, since the marginal cost curve does not change when W_1 change, the mapping from W_1 to x_2 will be unique[1]. Provided that the marginal utility of producing opium is greater than the marginal cost at zero production, this will be sufficient to ensure existence (however, not necessarily within a reasonable range of x_t). ■

$J(2, W_1)$ will be an increasing function in W_1: $\partial J(2, W_1)/\partial W_1 = u'(\cdot) > 0$ by the envelope theorem[2]. $J(2, W_1)$ will hence be increasing and concave in W_1.

Going to period $t = 1$, using the optimal value function for $t = 2$, we get the following

[1]Strictness in concavity is key for the result. Without this property on the utility function, the marginal utility curve may be horizontal and we may then not get a (unique) mapping from wealth to opium production.

[2]$\partial J(2, W_1)/\partial W_1 = \partial(u(h(W_1) + W_1) - g(h(W_1)))/\partial W_1 = u'(h(W_1) + W_1)(h'(W_1) + 1) - g'(h(W_1))h'(W_1) = [u'(h(W_1) + W_1) - g'(h(W_1))]h'(W_1) + u'(h(W_1) + W_1) = u'(h(W_1) + W_1) > 0$ since the first order condition for optimality makes the expression in the square brackets equal to zero.

maximization problem:

$$J(1, W_0) = \max_{c_1, x_1} \{u(c_1) - g(x_1) + J(2, W_1)\}$$

$$\text{s.t.}$$

$$W_1 = \begin{cases} x_1 + W_0 - c_1 & \text{if } x_1 + W_0 - c_1 \geq 0 \\ (x_1 + W_0 - c_1)\alpha & \text{otherwise} \end{cases} \tag{7}$$

$$c_1 \geq 0$$

$$W_0 \text{ given}$$

Due to the non–linear budget constraint, the solution to this problem will be somewhat involved. The reason is that the behavior of the farmer will be different for small values of wealth relative to small values of debt. For small values of wealth the consumption and production will be equal in the two periods, as wealth can be transferred without loss between the periods. Equal production and consumption in the two periods will be optimal since utility is concave. Small values of debt however, cannot be transferred without loss between periods, the reason being the salaam debt system increases the amount of debt with a factor of α if it is rolled over to the next period. For small values of debt it will hence be optimal to repay the entire debt now rather than to transfer it to next period and paying the salaam cost. Consumption and production will here not be equal in the two periods. A third case arises if the initial debt is large. For this level of debt it will no longer be optimal/possible to pay back the entire debt in one period. The reason is the extremely low consumption and the extremely high production level that is implied by choosing this strategy. It will hence be better to pay the salaam cost of borrowing, than to reduce consumption in the first period that much. There will then exist a level of debt for which the optimal strategy of the farmer switches from repaying everything in one period, to paying the salaam cost of borrowing. I label this level of debt W^*, the *salaam threshold.*

The solution to the problem is characterized in Proposition 2:

Proposition 2 *Assume that $J(2, W_1)$ is the optimal solution to the problem at $t = 2$, as identified in Proposition 1. Then there will be three different solutions to the problem at $t = 1$, depending on initial wealth W_0:*

A. $W_0 \geq 0, W_0 + x_1 - c_1 \geq 0$. The free optimization problem. Half of the initial wealth will be spend in each period; consumption and production will be equal in each period.

B. $W_0 < 0, W_0 + x_1 - c_1 < 0$. Debt is rolled over from $t = 1$ to $t = 2$, $W_0 < W^*$.

C. $W_0 < 0, W_0 + x_1 - c_1 \geq 0$. Debt is manageable in one period, $W_0 \geq W^*$.

We describe the cases:

A Given the shape of the utility function and the assumption that the discount factor $\beta = 1$, the intertemporal rate of substitution will be 1:

$$u'(c_1) + h'(W_1)[-u'(c_2) + g'(x_2)] - u'(c_2) = 0 \iff \frac{u'(c_2)}{u'(c_1)} = 1 \tag{8}$$

$$-g'(x_1) + u'(c_2)[h'(W_1) + 1] - g'(x_2)h'(W_1) = 0$$

$$\iff g'(x_1) = u'(c_2) = u'(c_1) \tag{9}$$

Consumption and production will be equal in both periods, and the production level is given by Proposition 1.

B We substitute in for the budget constraint, and get the following first order conditions:

$$u'(c_1) + \alpha h'(W_1)[-u'(c_2) + g'(x_2)] - u'(c_2)\alpha = 0$$

$$\iff \frac{u'(c_1)}{u'(c_2)} = \alpha \tag{10}$$

$$-g'(x_1) + u'(c_2)[h'(W_1) + W_1] - g'(x_2)h'(W_1)\alpha = 0$$

$$\iff g'(x_1) = u'(c_2)\alpha = u'(c_1) \tag{11}$$

Production will be higher and consumption lower in the first period relative to the second period.

C It will, for some level of debt $W_0 \in (W^*, 0)$, be optimal to repay the entire debt at $t = 1$ rather than to roll over parts of the debt. The solution to the problem at $t = 1$ will be then be given by Proposition 1:

$$g'(x_1) = u'(x_1 + W_0) \tag{12}$$

Production and consumption at $t = 2$ will be given by Proposition 1 with $W_1 = 0$. When will this strategy be chosen? Let x_2^* be the solution to the maximization problem at $t = 2$ for $W_1 = 0$. The "shadow price" of transferring debt to the next period at $t = 1$ is then $\alpha[u'(x_2^*) - g'(x_2^*)]$. Hence the farmer will repay the entire debt now if and only if the cost of producing opium now is smaller than the gain from not postponing it to the next period:

$$u'(x_1 + W_0) - g'(x_1) \leq \alpha[u'(x_2^*) - g'(x_2^*)]$$

Finally, W^* will be the W_0 that solves

$$u'(x_1 + W_0) - g'(x_1) = \alpha[u'(x_2^*) - g'(x_2^*)]$$

i.e. the level of debt that makes the farmer indifferent between choosing strategy B or C.

We have five endogenous variables $\{c_1, c_2, x_1, x_2, W_1\}$, hence we need five independent equations in each of the cases to solve the entire problem. Do we have that here? For all cases, c_2 will be determined by the last budget constraint with $W_2 = 0$, x_2 will be determined by (6), and W_1 will be determined by the intertemporal budget constraint (7). For case A, c_1 will be determined by (8) and x_1 by (9). For case B, c_1 is determined by (10) and x_1 is determined by (11). For the last case, case C, c_1 is found by solving (7) with $W_1 = 0$ for c_2, and finally x_1 is determined by (12).

We hence have shown that optimal opium production will be an increasing function of the level of debt, and that if the initial debt is high enough, the initial debt may influence the cropping strategy multiple periods ahead. In addition, opium farmers will find it optimal to shift production of opium to the present period relative to the next if they have debt, while not if they have positive wealth.

We have still not employed the minimum consumption constraints in the analysis. These constraints can be given an interesting interpretation. If we let the utility function and the production of opium be defined *relative* to producing other agricultural goods, we can let $\bar{c}$ denote the level of consumption that the household obtain if they do not produce opium. We can then change $\bar{c}$ according to changes in prices of other agricultural goods, and thereby introduce cropping alternatives into the model in a very simple way. The choice of opium production will then be non–continuous around the point where the consumption from producing opium goes below $\bar{c}$. If the set $S = \emptyset$, this will then simply mean that opium production is not worthwhile for the household. Hence opium will sometimes be more profitable than other crops, but farmers will be more prone to produce opium given that they have initial debt.

From this it follows that opium production in both periods will be functions of the level of debt, the salaam parameter α, and price incentives: $x_t = h(W_t, \alpha, \boldsymbol{p}_t)$.

We have modelled the decisions of the farmers in an extremely simplistic way, abstracting from uncertainty and moral costs. In addition we have imposed an end at time $t = 2$, which obviously governs the amount of opium produced for a given level of wealth. We will now introduce uncertainty into the model, and at the same time let the time horizon go to infinity, such that the effect of the imposed stop time will disappear. Moral costs will be introduced in Section 5.

3.2 Infinite horizon model with uncertainty

To be able to solve the infinite horizon model with uncertainty, we will have to make some further assumptions. Let $\{x_L, x_H\}$ be the two possible harvest levels, where $x_H > x_L$.

Expected production is then

$$Ex = ex_H + (1-e)\,x_L$$

where e is effort, $0 \le e \le 1$. We here assume that the probability of the good outcome is equal to the effort. This we can do without loss of generality, as long as the probability of the good outcome is an increasing and concave function of effort, our results still follow.

Higher effort implies higher production. The cost of effort is $g(e)$ where $g'(\cdot) > 0$, $g''(\cdot) > 0$ and $g'(1) = \infty$. In addition we assume that $u'(\bar{c}) = \infty$, to avoid solutions involving consumption at or below the minimum level. Finally we assume that consumption (and hence saving) is chosen ex post, while the production is decided upon ex ante.

We also need a no–Ponzi game condition, to avoid that the farmer roll over the debt infinitely. The no–Ponzi game condition will here be very simple, $W_t \ge x_{\max}\,\forall\, t$, i.e. the debt have to be below some exogenous given maximum level.

Let the intertemporal budget constraint be given by

$$W_t = k \cdot (x_t + c_t - W_{t-1})$$

and

$$k = \begin{cases} 1 & \text{if } x_t + c_t - W_{t-1} \ge 0 \\ \alpha & \text{if } x_t + c_t - W_{t-1} < 0 \end{cases}$$

Let W_0 be initial wealth. The Bellman equation will then be

$$\begin{aligned} V(W_{t-1}) &= \max_{c_t, e_t}\{u(c_t) - g(e_t) + E_t V(W_t)\} \\ &= \max\{u(c_t) - g(e_t) + E_t V[k(W_{t-1} + x - c_t)]\} \\ &= \max\{u(c_t) - g(e_t) + e_t V[k(W_{t-1} + x_H - c_t)] + (1-e_t)V[k(W_{t-1} + x_L - c_t)]\} \end{aligned}$$

Consumption is chosen ex post, while effort is decided upon ex ante. If x_t^H is realized, consumption is given by

$$u'(c_t^H) = V'\left(W_{t-1} + x_H - c_t^H\right) \tag{13}$$

I assume that $W_{t-1} + x_H - c_t^H \ge 0$ if the good state is realized. If the bad state x_L is realized, we will again have to consider cases, depending the level of the start–of–period debt. The cases will be approximately the same as the ones we identified in the previous section, hence I use the same labels.

A If $u'(c_t) = V'(W_{t-1} + x_L - c_t^L)$ has a solution where $W_{t-1} + x_L - c_t^L \ge 0$, this level of consumption will be chosen.

B If $W_{t-1} < W^*$, consumption will be given by $u'(c_t) = \alpha V'[\alpha(W_{t-1} + x_L - c_t^L)]$.

C If $W_{t-1} \in (W^*, 0)$, it will be optimal to consume less and not borrow. Consumption will be given by the accounting identity $c_t^L = W_{t-1} + x_L$. This strategy will be used when $u'(W_{t-1} + x_L) \leq \alpha V'(0)$, i.e. when the increased utility from consumption today is not enough to cover the costs of borrowing.

Solving the problem backwards, we then find the optimal effort level (remember that consumption now will be different in the two states, hence the expectation have to be taken over both the value function and the consumption in the present period):

$$g'(e) = u\left(c_t^H\right) - u\left(c_t^L\right) + V\left[k(W_{t-1} + x_H - c_t^H)\right] - V\left[k(W_{t-1} + x_L - c_t^L)\right]$$

To interpret this first order condition for effort, we have to look separately at the cases A, B, and C. For notational simplicity I introduce $q_t^i = W_{t-1} + x_i - c_t^i$ for $i = L, H$.

A The first order condition becomes

$$g'(e) = u\left(c_t^H\right) - u\left(c_t^L\right) + V\left(q_t^H\right) - V\left(q_t^L\right)$$

From this we find that effort is increasing in debt:

$$\begin{aligned} g''(e_t)\frac{\partial e_t}{\partial W_{t-1}} &= u'(c_t^H)\frac{\partial c_t^H}{\partial W_{t-1}} - u'(c_t^L)\frac{\partial c_t^L}{\partial W_{t-1}} + V'(q_t^H)\left(1 - \frac{\partial c_t^H}{\partial W_{t-1}}\right) - V'(q_t^L)\left(1 - \frac{\partial c_t^L}{\partial W_{t-1}}\right) \\ \iff \frac{\partial e_t}{\partial W_{t-1}} &= \frac{u'(c_t^H) - u'(c_t^L)}{g''(e_t)} < 0 \end{aligned}$$

since $u(\cdot)$ is assumed to be *strictly* concave. I have here used the ex post optimality condition for the savings/consumption choices from (13) for the good outcome and the optimality condition for case A for the bad outcome above, to substitute in for the value function.

B For case B, the level of debt is higher than the *salaam threshold* W^*, and some debt will hence be transferred to the next period. The first order condition for effort will be

$$g'(e) = u\left(c_t^H\right) - u\left(c_t^L\right) + V\left(q_t^H\right) - V\left(\alpha q_t^L\right)$$

Again we find that effort is increasing in debt:

$$\begin{aligned} g''(e_t)\frac{\partial e_t}{\partial W_{t-1}} &= u'(c_t^H)\frac{\partial c_t^H}{\partial W_{t-1}} - u'(c_t^L)\frac{\partial c_t^L}{\partial W_{t-1}} + V'\left(q_t^H\right)\left(1 - \frac{\partial c_t^H}{\partial W_{t-1}}\right) - \alpha V'\left(\alpha q_t^L\right)\left(1 - \frac{\partial c_t^L}{\partial W_{t-1}}\right) \\ \iff \frac{\partial e_t}{\partial W_{t-1}} &= \frac{u'(c_t^H) - u'(c_t^L)}{g''(e_t)} < 0 \end{aligned}$$

I have also here substituted in for the value function using the ex post optimality conditions (13) and case B.

C Finally we look at the last case, where debt is above the *salaam threshold*, i.e. all debt will be repaid in this period. The first order condition for effort becomes

$$g'(e) = u(c_t^H) - u(W_t + x_L) + V(q_t^H) - V(0)$$

Also here effort is increasing in debt:

$$g''(e_t)\frac{\partial e_t}{\partial W_{t-1}} = u'(c_t^H)\frac{\partial c_t^H}{\partial W_{t-1}} - u'(c_t^L) + V'(q_t^H)\left(1 - \frac{\partial c_t^H}{\partial W_{t-1}}\right)$$
$$\iff \frac{\partial e_t}{\partial W_{t-1}} = \frac{u'(c_t^H) - u'(c_t^L)}{g''(e_t)} < 0$$

Hence for all cases we find that optimal effort is strictly increasing in the amount of debt (or decreasing in wealth). The effort function $e(W_{t-1}, \alpha)$ will also be increasing in α, but I do not show that here.

In this and the previous section we have shown that optimal opium production (or effort devoted to opium production) will be strictly increasing in the level of debt. We have also shown that if we in the two–period model let the the minimum consumption level denote the consumption obtained by other agricultural production than opium production, and hence let our model be defined relative to other production, prices of other agricultural goods will influence the cropping choice. Optimal opium production will hence be a function $x_t = h(W_{t-1}, \alpha, \boldsymbol{p}_t)$, where $h(\cdot)$ is decreasing in the first argument and increasing in the second argument (α). The functional form of $h(\cdot)$ is however not identified, and as we can infer from Section 3.2, it will depend on the functional form of the cost function and utility function. In general it will be a nonlinear function of its arguments.

3.3 A model with moral costs

We have not yet discussed the costs that arise from breaking moral/religious norms when producing opium. In the two previous sections there was a simple switch from producing opium to not producing opium, the effort devoted to opium production was a choice of effort from a continuous set. There is however good reasons to assume that there are fixed costs from starting to produce opium which are not considered above, costs that relates to breaking moral and religious norms when producing opium, as opium production is considered to be haraam, or forbidden, under Islamic law.

To put this into the modelling framework presented above: the $V(W_{t-1})$–function represents the net present value (in utility terms) of choosing to produce an optimal level of opium given inherited wealth W_{t-1}. Assume that there also exists a value function that, for the same level of inherited wealth W_{t-1}, gives the net present value of a cropping strategy that does not involve opium production. Denote this value function $Q(W_{t-1})$.

We assume that $V(W_{t-1}) > Q(W_{t-1})$, i.e. if the farmer is non–religious, or if he does not care about the possible social stigma from producing opium, he would choose to act according to the optimal scheme characterized in section 3.2. But if there are moral costs to production of opium, a "fixed cost" from switching from other agricultural production to opium, this choice will not be that obvious. Denote these moral costs by m_i, where i indicates an individual farmer. The farmer's choice of cropping strategy will then be to choose opium production if and only if

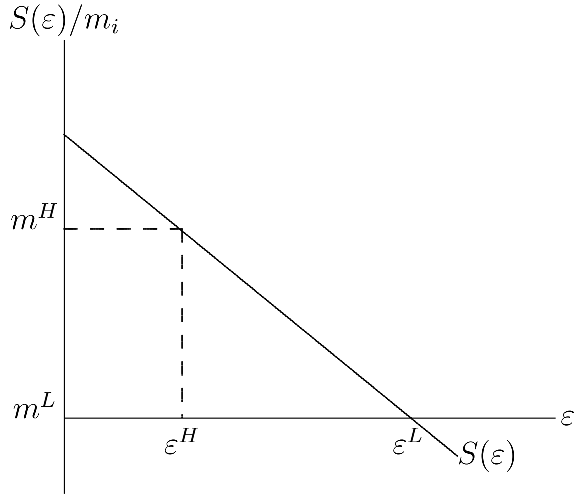
$$V(W_{t-1}) - m_i > Q(W_{t-1}) \tag{14}$$

Hence there will, for a farmer with a given moral cost term m_i, exist a threshold for debt for which the farmer switch from other agricultural production to opium production. But when this choice first is made, the moral cost term will disappear from consideration, and we will be back in the case discussed in section 3.2 where optimal effort is chosen as a function of inherited wealth. Why? Because if opium is chosen as a part of the farmer's cropping strategy, the fixed costs will have turned to sunk costs, as the moral stigma that is induced by opium production is not linked to the amount of opium produced. Hence for a farmer with a high moral cost term, when he first starts to produce opium he is likely to produce a lot, possibly utilizing all his available resources to opium production. The farmer may have accumulated debt over several years, and then finally, when the level of debt passes the crucial threshold in (14), the farmer changes strategy from other agricultural production to opium production. For these farmers we say that opium production is debt–induced.

I will make the last argument more clear using an example. The intuition behind the example is the following: the farmers that have a high moral cost term, we label them the "moralists", will for a sufficiently high level of debt start to produce opium. But, counter to the farmers that produce opium for profit, whom we label the "opportunists", the moralists will when first started to produce opium, produce at a much higher level. Their opium production can hence be said to be truly debt–induced, and they will have the characteristic that they produce closer to maximum capacity.

The example is the following. Assume that the farmer starts out with $W_{t-1} = 0$ and that the farmer has a consumption level of c. His expected pay–off will be $u(c) + Q(0)$ if no shock is to occur (remember that we for simplicity have set the discount factor to unity). Let us introduce some furter assumptions. First we restrict the values that the moral cost term can take, $m_i \in \{m_L, m_H\}$, where $m_L < m_H$. The two groups of farmers hence have their own moral cost term, the opportunists have m_L while the moralists have m_H. We further assume that the farmer is exposed to a random shock, ε, for which the level of his ex post consumption will be εc. The shock has a cumulative distribution function F. When the shock is realized, the farmer has two possibilities. The first is to

Figure 3: Why the "moralists" will produce higher amounts than the "opportunists"



continue to produce other agricultural produce than opium, to fix ideas let us assume that he produces wheat. The other possibility is to enter the credit market, but this he can only do if he is an opium producer, as opium is the preferred collateral for the lenders. He will then choose to produce opium if and only if

$$u(\varepsilon c) + Q < u(c) + V(c - \varepsilon c) - m_i$$

If he is to take up a loan, he borrows so that he can continue to have the consumption level c, i.e. he borrows $c - \varepsilon c$. The moral cost term occurs only in the first period, i.e. it is not part of the value function. The reason for this is that these costs are sunk cost when the farmer has started to produce opium, he has already been exposed to the social stigma. Rearranging we get that the farmer will choose to produce opium if an only if

$$m_i < V(c - \varepsilon c) + u(c) - u(\varepsilon c) - Q \equiv S(\varepsilon) \tag{15}$$

$S(\varepsilon)$ will be decreasing in ε (remember, small values of ε are the worst, as it affects consumption in a multiplicative way). From this it will be clear that, if all shocks are drawn independently from F, the average level of production for the moralists will be much higher than the average level for the opportunists. To see why, we draw a figure.

Figure 3 shows that the moralists need a much worse shock before they start to produce opium than the opportunists, i.e. before the inequality in (15) will be fulfilled (here m^L is set to 0). A fraction $F(\varepsilon^L)$ of the opportunists will produce opium, while only a fraction $F(\varepsilon^H)$ of the moralists chooses opium, where F is the cumulative density function of the shock ε. This means that the average production of the moralists will be much higher than the average production of the opportunists. Why? As shown in sections 3.1–3.2, opium production will be increasing in debt (here $c - \varepsilon c$). And since the moralists, on average, start to produce opium when they are worse off than the opportunists, they will have higher production levels than the opportunists.

The model with moral costs, and especially this last example, will be utilized in section 5, where the empiricial strategy is introduced.

4 Description of the data

The data set we use is obtained from United Nations, Office on Drugs and Crime (UNODC), and it is thoroughly described in UNODC (2004). In this section we will first comment on the methods used when collecting the data, then describe the most important variables for our analysis, and finally give some descriptive statistics and cross–tabulation of these variables. The empirical analysis is postponed to the next part of the thesis.

The data was collect by UNODC in October 2003, using local surveyors. The sampling frame was 13,980 villages (out of 30,706 known villages in the AIMS database), and their sampling strategy was to sample 2.5% of the sampling frame, i.e. 347 villages. However, due to security reasons, several districts in the parts of the Southern region could not be surveyed, and the UNODC ended up with collecting data from 308 villages.

An obvious problem with the data is that the sampling frame is *not* the entire Afghanistan; the UNODC have sampled farmers from provinces previously known to be producers of opium. Hence, any conclusion we can draw from our analysis cannot be said to be valid for entire Afghanistan, we must limit ourselves to give statements that (hopefully) are valid for the provinces where the sample is drawn from. In addition, the UNODC questions the method used to sample individual farmers from the randomly selected villages:

> "The debriefing sessions [...] revealed two shortcomings in the random selection process that could lead to an over–representation of the opium growing farmers in the sample: the first one is that in some villages the surveyors were directed by the headmen to interview farmers who were known in the village to have some experience in poppy growing. The second one is that in some cases, the surveyors, once in a village, were looking for farmers working in the fields. In many parts of eastern and southern Afghanistan, October was the time when farmers were preparing their fields for sowing opium poppy later in November." (UNODC, 2004, page 50)

Hence, given the sampling strategy and the shortcomings in the methods used for sampling of individual farmers, we will have to refrain from making claims involving aggregation based on the information found in this data set. However, as we are more interested in the motivation of the farmers rather than to make aggregate predictions, these shortcomings in the sampling strategy will not be of major importance for us. As long as the sampled farmers that are producing opium are representative for the population of opium producing farmers (and the non–opium producing farmers are representative for the population of non–opium producing farmers), our estimates will remain unbiased. We have no reason to doubt that this is the case.

The data set consists of two parts: one part where 922 farmers are randomly selected from 308 villages in opium producing provinces, and another part consisting of interviews with the heads of the same 308 villages. It has proved difficult to get hold of the questionnaire for the survey, however, it has been possible to infer what the questions are likely to have been from the data file. The two parts of the data set have been merged and collapsed so that it is fit for quantitative analysis using Stata.

Important descriptive statistics are given in Table 1[3]. We see that there is a huge increase in the number of farmers producing opium in 2004 relative to 2003. It is important to note that the survey was conducted in October 2003, so the figures for 2004 are the farmer's cropping *intentions* for 2004, hence they do not necessarily correspond with the actual outcome. Farmers may have had an incentive to report too high figures of opium production in 2004 to try to get attention and support from the government and local authorities. However, as the increase in production from 2003 to 2004 in the whole Afghanistan ex post has proven to be considerable (see Figure 1 for the increase in hectares of land devoted to opium production), we do not consider this to be a serious threat to the validity of our study.

Table 1: Descriptive statistics

	% of farmers	Observations
Opium 2003	.76	922
Opium 2004	.87	922
Monocrop opium 2003	.09	922
Monocrop opium 2004	.18	922
Start before 2001*	.63	804
Start 2002*	.10	804
Start 2003*	.19	804
Start 2004*	.08	804

* % of farmers producing opium in 2004

In table 2 and table 3 we report cross–tabulations of opium producers and debt (relative figures are given in parentheses). Debt is here a dummy variable taking the value of 1 if the farmer has taken up debt prior to 2003/in 2003 and 0 otherwise. Opium producer is also a dummy variable taking the value of 1 if the farmer is an opium producer and 0 otherwise. We restrict the sample to include only those respondents who are farmers in 2003 and 2004 respectively. In table 4 we exclude debt to family members, as a robustness test. The impression is still the same, debt seems to affect the cropping pattern of the farmers.

The impression we get from these cross–tabulations is that the conditional probability of being an opium producer is greater if the farmer has debt, relative to not having debt.

[3]The amount of opium producing farmers is different from what UNODC (2004) reports. UNODC (2004) may have added information to the data set that is not publicly available, but as we do not have this information, we use the data set as it is provided to us by the UNODC.

We investigate whether this is the case using t–tests for difference of means:

$$t = \frac{\bar{x} - \bar{y} - (\mu_x - \mu_y)}{s_p\sqrt{1/n + 1/m}} \sim t_{df} \tag{16}$$

where $\bar{x}$ denotes average of opium producing farmers given debt > 0, $\bar{y}$ denotes average of opium producing farmers without debt, m denotes the number of farmers that do not have debt, n denotes the number of farmer that have debt, and finally

$$s_p = \sqrt{\frac{(n-1)S_x^2 + (m-1)S_y^2}{m+n-2}}$$

denotes the pooled sample variance. The t–statistic is t–distributed with $(m+n-2)$ degrees of freedom. We are interested in testing whether the mean is different in the two subsamples, hence we apply a two–sided test:

$$h_0 : \mu_x = \mu_y, h_A : \mu_x \neq \mu_y$$

We use a significance level of 1 %.

We first test whether there is a difference in the cropping pattern in 2003 for the farmers with outstanding debt from before 2003 relative to farmers without outstanding debt. The data for the test can be found in Table 2[4]. The test–statistic (16) becomes

$$\hat{t} = \frac{.84 - .75}{\sqrt{.17(1/233 + 1/671)}} \approx 2.617$$

The critical value is $t^*_{\alpha/2} = 2.576$ with 902 degrees of freedom, and we can hence reject the null hypothesis that the cropping pattern is similar for both groups of farmers.

Table 2: Cross–tabulation, debt and opium production 2003

	Loan < 2003		
Opium 03	0	1	Total
0	165 (.25)	38 (.16)	203
1	506 (.75)	195 (.84)	701
Total	671 (1.0)	233 (1.0)	904

Next we test whether the cropping pattern in 2004 depends on whether the farmer has taken up a loan in 2003. The reason that we include this test as well as the former, is twofold. The first reason is that we have more data on debt taken up in 2003 than debt taken up earlier than 2003; we have data on who the creditor is. This allows us to do some more robustness tests, for example we can test whether excluding debt to family still produces the same result. Second, this allows us to test the direct effect of debt taken

[4]Standard deviations are not provided in the tables, but are available on request.

up one year in advance of the production, while the outstanding debt from before 2003 variable does not specify which year the debt is from. It is reasonable to assume that if a farmer has taken up debt in e.g. 1998, and the loan did not induce opium production at that time, it is unlikely to have any other effect several years later. Hence, debt taken up one year in advance gives us a more direct test of our hypothesis.

The test statistic (16) becomes

$$\hat{t} = \frac{.92 - .85}{\sqrt{.10(1/485 + 1/424))}} \approx 3.55$$

The null hypothesis is rejected, even clearer than in the previous test. Again this seems to support our general hypothesis, namely that opium production is partly debt–induced. In addition, it seems to be that last year's debt is a stronger determinant of opium production, relative to debt taken up earlier than one year in advance of the harvesting.

Table 3: Cross–tabulation, debt and opium production 2004

	Loan in 2003		
Opium 04	0	1	Total
0	73 (.15)	32 (.08)	105
1	412 (.85)	392 (.92)	804
Total	485 (1.0)	424 (1.0)	909

The last t–test is a robustness test. Here debt to family members is excluded, and we hence look at whether the cropping profile is different for farmers that have debt to external creditors relative to debt–free farmers. A priori we expect to find that the effect still is significant, however we expect to find a lower t–statistic than the previous test since the variable debt is a superset of debt without debt to family members.

The test statistic (16) is

$$\hat{t} = \frac{.92 - .85}{\sqrt{.10(1/513 + 1/396)}} \approx 3.31$$

Again, the test rejects the null hypothesis.

Table 4: Cross–tabulation, debt (not to family) and opium production 2004

	Loan in 2003 (not fam.)		
Opium 04	0	1	Total
0	75 (.15)	30 (.08)	105
1	438 (.85)	366 (.92)	804
Total	513 (1.0)	396 (1.0)	909

This section has shown that the probability of being an opium producing farmer, conditional on debt, is greater than the probability of being an opium producing farmer, conditional on not having debt. The findings are consistent with our main hypothesis, namely that opium production is partly debt–induced. However, doing simple correlation analysis as the ones we have just done, may be misleading. One of the pitfalls is that the dependent variable and the independent variables are both correlated with a third (latent) variable, and hence that the correlation we observe is not causal, it is just a product of (perhaps causal) correlation with a third (unknown) variable. To remedy this, we perform more thorough tests in the next section, controlling for other possibly related explanatory variables.

5 Empirical strategy and results

Our key equation from the theoretical model states that opium production in year t is a function of the level of debt taken up at $t-1$ and the stock of outstanding debt from the years $0, \ldots, t-2$. In addition the salaam parameter, α, is important, along with price incentives. Price incentives where introduced into the model by letting the minimum consumption level $\bar{c}$ represent the consumption obtained from agricultural production without producing opium, and the model of opium production where then defined *relative* to the production of other agricultural goods. A modified version of the key equation is therefore

$$x_t = h\left(d_{t-1}, \sum_{s=0}^{t-2} d_s, \alpha, \boldsymbol{p}_t\right) \tag{17}$$

where d_t is the amount of debt taken up at time t ($= -(W_t - W_{t-1})$), using the notation from the theoretical part).

We started considering the moral costs that arise from breaking the moral/religious norms when producing opium in section 3.3. As stated, opium and opium production is considered to be haraam, or forbidden, under Islamic law, and the farmers that are producing opium are hence deliberately violating this code of conduct. It is reasonable to assume that violating these moral norms will induce a "cost" for two reasons. First, the farmers are producing something that they themselves consider to be illegal, and, depending on the strength of their own religious belief, this induces a subjective moral burden. Second, opium production may have an impact on the farmer's reputation in the local community, and it may even lead to social sanctions. We will hereafter refer to these costs as moral or religious costs.

Equation (17) is a deterministic function governing the amount of opium production for a given amount of initial debt, abstracting from considerations like moral and religious norms. Introducing these moral costs will, since farmers are assumed to be heterogeneous with respect to the strength of their own religious belief, give rise to heterogeneity in the amount of produced opium for given initial debt.

Religious sentiments are not measurable, at least they are not explicitly considered in the data material we have obtained. We treat the religious costs as unobservables, i.e. as residuals in the multinomial logistic regressions. Each individual farmer is assumed to know their own moral cost term, but from a modelling perspective, as we do not observe the individual's moral cost term, we consider it to be drawn individually and independently by the farmers from some probability distribution.

We add this unobservable part of the utility to the deterministic part given in (17), and get the following equation

$$x_{i,t} = h\left(d_{t-1}, \sum_{s=0}^{t-2} d_s, \alpha, \boldsymbol{p}_t\right) + \epsilon_{i,t}$$

As the specification stands now, this seems to be a perfect setting for an OLS regression. However, several arguments may be put forward for why this not a good approach. Firstly, in the theoretical part we showed that the h–function may be highly nonlinear, making a linear (in variables) regression approach inappropriate. Secondly, we do not have any data on wealth, only debt is registered in our data set, hence parts of some of the most important explanatory variables are unobserved (coded as 0). Thirdly, as negative production of opium is not a possibility, the dependent variable is censored at 0. This makes the distribution of the error term truncated, and the OLS estimates will be biased as the error term will have a non–zero expectation for certain values of the explanatory variables. The third problem we could remedy using Tobit estimation. Unfortunately, this model does a poor job in explaining the data, possibly because of nonlinearities in the h–function or the missing parts of the independent variables (the pseudo-R^2 is about .07 in all regressions). In addition, the Tobit method imposes extremely strong assumptions on the distribution of the error terms (Cameron and Trivedi, 2005, chapter 16). Results from the Tobit regressions are given in Table 6.

Another reason why we do not get a good fit may be that moral costs may introduce fixed costs from switching from production of some other agricultural crop to opium, as considered in Section 3.3. This is related to the argument that the dependent variable is truncated at 0; for many values of the moral cost term the farmers would like to produce negative amounts of opium. However, as this is not possible, it is coded as 0. Hence it might be more interesting to look specifically at the switch from producing opium to not producing opium, which is what we turn to now.

The approach we choose is to estimate choice probabilities of producing opium using discrete choice models. The dependent variable will here be an indicator function for whether the farmer is producing opium or not. In this way we remedy the problems that arise from the possibly nonlinear h–function. The problem of a censored dependent variable will also be solved, the distribution of the error term will no longer be truncated, and we use a more direct approach to address the problems of fixed cost of starting to produce opium.

We first focus on the choice of opium production in general, where we look at the binary choice between producing and not producing opium. After that we look at the different groups of farmers, modelling the choices using multinomial logistic regressions.

5.1 Conditional probability of producing opium as a function of debt

We first test our model using a binary logit approach, where the dependent variable is an indicator function for whether the farmer is producing opium or not. We are interested in estimating choice probabilities for the two alternatives, given the characteristics of the farmers. To accomplish this we run multinomial logistic regressions, estimating alternative–specific parameters.

A formal specification of the approach is the following: let

$$y_t^i = \begin{cases} 1 \text{ if } i \text{ chooses to produce opium in year } t \\ 0 \text{ otherwise} \end{cases}$$

and let utility for the different alternatives be given by $U_{ij,t}$, where i refers to an individual farmer and $j \in \{0, 1\}$. As parts of $U_{ij,t}$ will, by necessity, be unobservable, we can split this utility into two parts

$$U_{ij,t} = V_{ij,t} + \epsilon_{ij,t} = \boldsymbol{x}'_{i,t}\boldsymbol{\beta}_j + \epsilon_{ij,t} \tag{18}$$

where $V_{ij,t}$ is the representative utility, $\epsilon_{ij,t}$ is a random error term, $\boldsymbol{x}_{i,t}$ is characteristics of the farmer and $\boldsymbol{\beta}_j$ is the alternative–specific parameters to be estimated[5]. The representative utility is specified to be linear in parameters. For notational simplicity, we drop the time subscript from now on.

The farmer chooses the alternative that gives him the greatest utility; he chooses j if and only if $U_{ij} > U_{ik} \, \forall \, k \neq j$. As we are in a probabilistic framework, ϵ_{ij} is unobserved, we will have to model choice probabilities rather that deterministic choices (the derivation is from Train (2003)):

$$\begin{aligned} P_{ij} = \Pr[i \text{ chooses } j] &= \Pr[U_{ij} > U_{ik} \, \forall \, k \neq j] \\ &= \Pr[V_{ij} + \epsilon_{ij} > V_{ik} + \epsilon_{ik} \, \forall \, k \neq j] \\ &= \Pr[\epsilon_{ik} - \epsilon_{ij} < V_{ij} - V_{ik} \, \forall \, k \neq j] \end{aligned}$$

From the equation we see that only differences in utility matters, only relative parameters can be identified. Here the estimated parameters will be relative to the parameters of the 0–alternative. To relate this to the theoretical part above, the choice probabilities for our model can be given by

$$\Pr[i \text{ chooses opium}] = \Pr[V(W_{t-1}) - m_i > Q(W_{t-1})]$$

[5] All vectors (denoted by bold type) are column vectors unless otherwise stated.

where $V(W_{t-1})$ is the value function defined in the theoretical part for given inherited debt W_{t-1}, m_i is the individual farmer's moral cost term, and $Q(W_{t.1})$ is the discounted value of a program where opium is not part of the cropping strategy, given the same inherited debt W_{t-1}. Here I have normalized the moral cost of producing other crops than opium to zero. In principle this term could be anything, as long as m_i is scaled appropriately.

The functional form of P_{ij} will obviously depend crucially on the assumed distribution of the error terms. Following the literature on discrete choice, we specify the error term (ϵ_{ij}) to be i.i.d. Extreme Value Type 1 (EV1) distributed. The choice of the extreme value distribution may at first sight seem somewhat arbitrary, but there exist good theoretical reasons as to why the error term should have exactly this distribution. If we believe that Luce (1959)'s independence of irrelevant alternatives (IIA) axiom holds for the choices in question, McFadden (1974) proved that the only random utility maximization model (RUM) that is consistent with the IIA axiom, is a RUM with extreme value distributed error terms. Hence we have given an axiomatic justification for the distribution of the error terms; if we believe that choices are made according to IIA, the error terms have to be extreme value distributed[6]. The distribution of the error term follows directly from behavioral assumptions, rather than some ad hoc hypothesis on the distribution of the error term as is often proposed when using an OLS regression approach. It is also interesting to note that the estimated coefficients are not very dependent on whether the specification of the distribution of the error term is EV1 or not, the difference in estimated coefficients between the binary logit and the binary probit (where the error term is assumed to be normally distributed), for example, are almost indistinguishable (Dagsvik, 2006).

We specify $\boldsymbol{x}_n$ to contain the following characteristics (a priori expectations on the signs are given in parantheses):

Loan in 2003 (+) Amount of loan taken up in 2003. In the data set three different currencies are used; Afghani, Pakistani Rupee (.850) and Iranian Toman (.584). We have converted the loans into Afghani using the exchange rates given in parentheses. The rates are the official exchange rates at January 1., 2004 (cash and transfer average), provided by the Afghan Central Bank. An obvious flaw is that the exchange rates are likely to vary much at different places in Afghanistan, depending on distance to banks etc. However, as we do not have any data on this, we must resort to using official rates. In addition we would like to have exchange rates from

[6] The IIA property does not have any meaning when there are only two choices, as here. However, as we later will use multinomial logit, we introduce the terminology already here. The reason that the IIA property does not make sense when there are only two alternatives is that there does not exist any subset of the choice universe on which the agent can make a choice. Any strict subset of the choice universe will necessarily contain only one element, and hence there will not be any choice made by the agent. The IIA choice axiom states that the choice probabilities when calculated on any choice set that is a strict subset of the entire choice universe should not be dependent on the choices not contained in the subset in question, more formally: $\Pr_i(S) = \Pr_i(B)\Pr_B(S)$, where $i \in B \subset S$.

October 2003 (the month when the survey was undertaken), however we have not been able to obtain these. The variable is scaled down by a factor of 1000 to obtain more easily interpretable estimates.

Outstanding loan in 2003 (+) The amount of outstanding loan taken up before 2003, converted to Afghani using the exchange rates given above. Scaled down by a factor of 1000.

Farmgate price of opium (+) Farmgate price of opium in the 2003 season. The variable is calculated in the following way:

1. If the farmer has reported a price, this price is used.
2. If the farmer has not reported a price, the average of the reported prices in the village is used.
3. If none of farmers in the village have reported a price, the average in the district is used.
4. If none of the farmers in the district have reported a price, the average in the area is used.
5. If none of the farmers in the area have reported a price, the average in the region[7] is used.

If no price is calculated after going through this five–step algorithm, the price variable is set to `missing`, indicating that opium is not an option for that particular farmer.

Farmgate price of wheat (+/−) The farmgate price for wheat in 2003. The variable is calculated using the same algorithm as for the opium price variable, but one new step is added at the end: if the variable is still missing after having used the average in the province, the average of all the prices is used. The reason for this is that wheat farming is always an option for the farmer, as wheat can be used for own consumption[8].

Farmgate price of vegetables (+/−) Farmgate price of vegetables in 2003. The variable is calculated in the same way as the wheat variable. Introduced into the utility for the same reason as the price of wheat, namely to catch opportunity cost of the land.

Price of labor (−) Calculated using the same method as for the prices of wheat and vegetables.

[7] village $<$ district $<$ area $<$ province $<$ region.

[8] It is reasonable to assume that opium is not produced for own consumption, at least not purely for own consumption.

Eradication in 2003 Variable indicating whether there has been eradication of opium crops in the village in the 2003 season. Unfortunately there are many missing observations here, so we again apply an algorithm to calculate eradication probabilities:

1. If the village headman have report a response, this response $\in \{0, 1\}$ is used.
2. If the variable is reported as missing, the average in the district is used.
3. If the variable is still reported as missing, the variable is coded as 0 (= no eradication).

Own in jeribs The amount of land owned by the farmer, introduced to capture to which social classes/wealth groups the opium producing farmers belong. A problem is that this variable might be endogenous to the decision we are analyzing, the reason being that the farmer may own land because they started to produce opium, while other may be landless because they refused to produce opium. Whether farmers choose to do this or not, will depend on their moral cost term, hence some rich farmers will be rich because they started to produce opium, some will be rich because they are rich and therefore need not produce opium, while the same is true for poor farmers: some are poor because they have always been poor, while some are poor because they have been forced to sell their land. The problem is hence the lack of time dimension in the data. As we only have cross–sectional data, we are bound to miss out on some important and interesting points regarding land ownership and opium production.

Regional dummies Dummy variables to catch region–specific fixed effects[9].

We include prices in 2003 as explanatory variables for production in 2004, i.e. we assume adaptive expectations. This is, as we see it, a natural assumption about the formation of expectations, since the planting is done one year in advance. A priori expectations on the sign of the coefficients are given in parentheses in the list of variables above. Clearly, we expect a high opium price to have a positive influence on the probability of producing opium. The farmgate prices of wheat and vegetables are included to take into account the opportunity cost of land. We do not have any a priori expectation on the signs of these two coefficients, the reason being that higher wheat/vegetable prices have two effects, a "substitution" effect and an "income" effect, that can draw in opposite directions. First, increased wheat price relative to opium will make the probability of opium production smaller (the substitution effect). Second, increased wheat price also means that it is harder to earn enough to buy food by any alternative production, and the farmers are forced to produce the most profitable crop, opium (the income effect). The

[9]There are five regions, the division of provinces into regions follow UNODC (2004). The reason why we do not include lower–level dummies (e.g. province or district specific dummies) is that introducing such dummies gives severe estimation problems (collinearity and perfect predictors).

eradication variable is included to control for the risk of eradication. It is obviously too simple to assume that the risk of eradication is proportional to the year before, however as the planting is done in 2003, recent eradication in the village may have influenced the decision on what to crop in 2004. Eradication may also have another effect; farmers that experienced eradication in 2003 may have problems repaying their debts due to the loss of their produce, and therefore roll over the debt to the next harvesting. But due to the salaam system, the amount of debt will then have increased by a factor of α, and they have to produce more opium to be able to repay the loan (discussed at length later).

The results are given in Table 5. We have run four logit regressions, the two first include the amount of loan taken up in 2003 as an explanatory variable, while the two last use loan taken up in 2003 excluding loans from family members. It is reasonable to assume that loan from family member have a different effect on opium production than loans from external creditors, hence we test both variables. We both include and exclude regional–specific dummies, regression (2) and (4) include regional–specific dummies.

Surprisingly, the amount of loan taken up in 2003 does not seem to have an impact on the probability of producing opium, not even when debt to family members is excluded. The estimated coefficient is positive as expected in all of the regressions, but the standard error is too big to make us able to reject the null hypothesis that the coefficient is equal to zero. The t-value is much closer to being significant in the case when debt to family members is excluded, but still not significant at a 5% level.

Price incentives seem to play an important role, the probability of producing opium depends negatively on the wage rate in the district and positively on the observed vegetable price, opium price and wheat price. As opium is very labor intensive, it is not surprising that a high wage rate in the area lowers the probability of being an opium producer. In addition, a high wage rate increases the opportunity cost of labor, making the use of own household labor to produce opium more expensive. The price of vegetables has a positive and highly significant coefficient. The reason for this may be, as stated, that an increase in the price of food stuff makes the farmers relatively worse off, and they must produce the most profitable crop to be able to buy food. The income effect is hence stronger than the substitution effect. The price of wheat is only significant when regional specific dummies are included, which may be taken as evidence for that there are regional–specific shifts in the wheat price, possibly due to that prices are determined in regional markets. The price of opium is only significant when regional dummies are excluded, this may be due to the algorithm we have used to replace missing observations on opium price. This last step of the algorithm replaces missing observations with the average of the regional prices, and this may create some collinearity between opium prices and the regional–specific dummies. The sign of the opium price coefficient is positive as expected in all specifications.

The amount of land owned is significant and positive in specification (1) and (3), and the estimated coefficients are almost equal between specification (1) and (2), and (3) and (4). Due to that the coefficient is almost equal across the specifications, and significant in two of them, we draw the tentative conclusion that opium farmers not necessarily are parts of the poorest segments of the population, as the conditional probability of producing opium depends positively on the amount of land owned.

The most surprising result, however, is that eradication in the area in 2003 contributes positively to the probability of being an opium producer in 2004. The effect is also significant at a 5% level when regional dummies are included in the logit regressions. As explained in the theoretical part, this may be because the farmer ends up in a debt trap if his crops are eradicated, and the farmer is hence forced to produce opium the next period. It is however important to note that we have not identified which of farmers that actually experienced the eradication, eradication here simply means that there has been eradication in the village or in the area surrounding the village in 2003. Another reason for that eradication increases the probability of opium production, is that eradication may free farmers to work for others, creating a boost in production since opium is extremely labor intensive.

5.2 Tobit regression

As stated in the previous section, we have run a Tobit model on the data. The results are presented in Table 6. Again we find that price incentives are important, both the price of opium and the price labor have a strong impact on the amount of opium produced. But counter to the previous results neither wheat prices nor vegetable prices are significant, and the signs have also changed. Eradication in 2003 also seems to be strongly influencing the amount of opium produced in 2004, if there has been eradication in the village the amount of opium produced is shifted up by approximately 1.5 jeribs. The claim that the opium producers are not from the poorest segments of the population is again verified, the amount of opium produced depends positively on the amount of land owned, the effect is significantly different from 0 and the estimated coefficients almost the same in all the regressions. Neither here nor in the binary logit case we find that debt influence the cropping choice, which is counter to widespread belief. But as stated, the Tobit regressions fits the data poorly, the pseudo-R^2 is about .07 in all the regressions (not reported in Table 6).

5.3 Discriminating between "moralists" and "opportunists"

From the results to the binary logit case, we find no significant effect of debt on the conditional probability of being an opium producer (see Table 5, and the discussion of the results in Section 5.1). In this section we modify this claim, by introducing more "types"

Table 5: Binary logit estimates

	(1)		(2)		(3)		(4)	
Expl.var.	Coef.		Coef.		Coef.		Coef.	
Own in jeribs	.019219	*	.019362		.018404	*	.01824	
	(.009)		(.010)		(.009)		(.010)	
Loan 03	.002795		.00129					
	(.002)		(.002)					
Loan 03, not fam.					.003358		.002463	
					(.002)		(.002)	
Outstanding loan 03	.008094		.003064		.005352		.00296	
	(.009)		(.005)		(.006)		(.005)	
Price of vegetables	.209465	***	.232461	***	.205245	**	.231087	***
	(.074)		(.070)		(.074)		(.070)	
Price of wheat	.097982		.260764	***	.102506		.26275	***
	(.065)		(.087)		(.065)		(.087)	
Price of labor	-.003962	*	-.002636	*	-.002998	***	-.002437	**
	(.002)		(.001)		(.001)		(.001)	
Price of opium	.000061	*	.000041		.000062	*	.00004	
	(.000)		(.000)		(.000)		(.000)	
Eradication in 03	.312373		.665975	*	.292798		.648433	*
	(.292)		(.305)		(.292)		(.305)	
Constant	-.412404		-1.788991		-.55089		-1.791932	
	(.659)		(1.019)		(.622)		(1.013)	
Regional dummies	No		Yes		No		Yes	
Log likelihood	-302.60		-281.94		-302.79		-281.88	
Iterations	7		7		6		6	
Obs	909		909		909		909	

* significant at 5%; ** significant at 1%; *** significant at .5%

of opium farmers.

Why does our model fail? We believe that one in general can say that there are two groups of opium producing farmers. The first group is the group that we are interested in, the poor farmers who are forced to produce opium to be able to obtain loans. However, there is also another group, the "opportunists" (introduced in section 3.3), who produce opium because it is the most valuable crop. The motivation for these two groups to produce opium is very different, and to measure the effect that the salaam system and debt have on the production of opium, we need to be able to distinguish between the two groups of farmers. One solution to this problem is to partition the sample of opium producing farmers into groups, where the selection criterion is the relative amount of opium produced, following the example in section 3.3. If we find that different factors are important to farmers in the different groups, and especially if debt is an important variable for the groups where the relative amount of opium produced is close to one while unimportant for the groups that produce little or no opium, this can be seen as evidence

Table 6: Tobit estimates

	(1)		(2)		(3)		(4)	
Own in jeribs	.171	**	.170	**	.170	**	.169	**
	(.009)		(.009)		(.009)		(.009)	
Outstanding loan 03	.002		.002		.001		.001	
	(.002)		(.002)		(.002)		(.002)	
Loan 03	.001		.001					
	(.000)		(.000)					
Loan 03, not fam.					.003		.004	
					(.004)		(.004)	
Price of vegetables	-.131		-.085		-.134		-.089	
	(.090)		(.091)		(.09)		(.091)	
Price of wheat	-.030		-.029		-.033		-.031	
	(.032)		(.032)		(.032)		(.032)	
Price of labor	-.011	**	-.013	**	-.011	**	-.013	**
	(.002)		(.002)		(.002)		(.002)	
Price of opium	.000	**	.000	**	.000	**	.000	**
	(.000)		(.000)		(.000)		(.000)	
Eradication in 03	1.594	**	1.406	**	1.570	**	1.381	**
	(.433)		(.438)		(.433)		(.438)	
Constant	1.041		1.532		1.009		1.473	
	(.866)		(1.404)		(.873)		(1.408)	
Regional dummies	No		Yes		No		Yes	
Observations	909		909		909		909	
# Left-censored observations	105		105		105		105	
Log likelihood	-2537.28		-2523.48		-2538.03		-2524.53	

Standard errors in parentheses
* significant at 5%; ** significant at 1%

that opium production is debt–induced for the farmers in the first category. The rationale for this partitioning is based on three different grounds. First is the extreme convexity in the costs of producing opium relative to other agricultural goods, we assume that the profit maximizing level of opium production is strictly smaller than using the entire amount of arable land available to a farmer to opium production. To produce one hectare of opium takes approximately 350 person days, while a hectare of wheat only takes about 40 person days (UNODC, 1999a). Since the Afghan farmer often attributes zero cost household labor, the alternative cost is zero since female household members often are prohibited from working in other parts of the economy than household production, there will be a kink in the cost curve at the point where the farmer needs to hire non–family labor. Second is an argument about the need for food security. Mansfield (2005) finds that food security deters households from exclusively cultivating one particular crop, if possible. Third, as we saw in Section 3.3, when the farmers have a high moral cost term, passing the threshold level of debt that induces opium production may lead to huge amounts of production, as the farmers have refrained from managing their debt the optimal way due to the moral costs of opium production. But when they first start to produce opium, the cost from the social stigma becomes sunk costs, as the moral costs are not related to the amount of opium produced. This we will utilize in the following to distinguish between the two groups of farmers. In the example in section 3.3, we found that moralists either produce nothing or very much, while the opportunists diversify their production. We impose the simplifying assumption that the moralists chooses a relative production level $\in \{0, 1\}$, while the opportunists choose a production level in the interval $(0, 1)$. To focus on these groups, and what the important determinants are for the different groups, we again apply a multinomial logistic regression approach, modelling alternative–specific parameters.

The alternative–specific parameters will be defined relative to the baseline alternative, which here is to not produce opium, since only relative parameters can be identified. The dependent variable is now a step function:

$$y_i = \begin{cases} 0 \text{ if } i \text{ is a farmer but does not produce opium in 2004} \\ 1 \text{ if } i \text{ have a mixed cropping strategy which also includes opium in 2004} \\ 2 \text{ if } i \text{ monocrops opium in 2004} \end{cases}$$

The derivation of choice probabilities follows directly from section 5.1, by letting j take values in $\{0, 1, 2\}$ instead of $\{0, 1\}$. We use exactly the same set of explanatory variables as in the binary logit case.

The results from the multinomial logit regressions can be found in Table 7. We have run two regressions, one including and one excluding regional–specific dummies. As we are estimating alternative–specific parameters, and the parameters are given relative to the 0–alternative, the table contains two columns for each regression. The first column gives the parameters for the farmers who produce opium on parts of their land, the "opportunists", and the second column contain the parameters for the group who devote all their land to opium production, the group for which we claim that opium production is debt–induced.

We first note that eradication in 2003 has a different effect on the two groups, which may explain the findings from the binary logit case. Eradication has a much stronger effect on the group that devote all their land to opium than the other, which can be seen as evidence for the story that farmers who experience eradication will have to increase their production even more the next year to be able to repay their debts. Also debt, excluding debt to family members, has a strong and positive effect for the farmers in category 2, supporting that opium production is debt–induced for these farmers: the conditional probability of ending up as a farmer who monocrops opium is strictly increasing in the amount of debt taken up in 2003, controlled for the amount of the land the household

Table 7: Multinomial logit estimates

	(1)				(2)			
	Polycrop		Monocrop		Polycrop		Monocrop	
Own in jeribs	.025015	*	-.105485	***	.025252	*	-.107302	***
	(.010)		(.023)		(.011)		(.024)	
Outstanding loan 03	.00449		.007364		.002309		.005015	
	(.006)		(.006)		(.005)		(.005)	
Loan 03, not fam.	.002679		.009009	**	.002001		.007404	*
	(.002)		(.003)		(.002)		(.003)	
Price of vegetables	.201512	**	.209862	*	.226104	***	.243306	***
	(.074)		(.083)		(.070)		(.081)	
Price of wheat	.074009		.19622	**	.239755	**	.362036	***
	(.068)		(.073)		(.087)		(.092)	
Price of labor	-.002881	***	-.002347		-.002411	***	-.001449	
	(.001)		(.002)		(.001)		(.002)	
Price of opium	.000057	*	.000095	***	.000051		-.000007	
	(.000)		(.000)		(.000)		(.000)	
Eradication in 03	.157325		.757856	*	.509381		1.0502	***
	(.297)		(.342)		(.309)		(.361)	
Constant	-.51147		-3.08238	***	-2.081076	*	-2.699812	*
	(.630)		(.805)		(1.027)		(1.229)	
Regional dummies	No				Yes			
Log likelihood	-634.49				-603.01			
Iterations	6				6			
Obs	909				909			

Standard errors in parentheses

* significant at 5%; ** significant at 1%; *** significant at .5%

owns. We also see that the amount of land affects negatively the probability of ending up as a category 2 farmer. This may be due to two reasons, one is that since production costs are highly convex for opium relative to other crops, more land will reduce the probability of the farmer devoting all his land to opium. The other reason that we get this result is that poor farmers who produce opium more often do it out of necessity, they devote all their arable land to production, while richer farmers produce at a level where they maximize profits.

Price incentives are important for both groups. The price of labor seems to influence only the choice of ending up in category 1 relative to not producing opium, i.e. if the price of labor goes up these farmers switch to producing other crops, while the effect on the category 2 farmers is much smaller. Again we can interpret this as evidence for the story that their production is debt–induced, as the supply of opium for this group is relatively inelastic to changes in wages.

5.3.1 Test of the implied correlation structure for the error terms

There is however one potential problem with the multinomial logit approach, namely the assumed independence of the error terms across alternatives. It might be reasonable to assume that the error terms for alternative 1 and 2 share some features and hence are correlated, as they both involve production of opium. This correlation structure may be captured in a nested logit model, and we estimate this version of the model to test whether the multinomial logit approach is appropriate, or whether we need to employ more sophisticated methods. In the nested logit approach, we can think of the choices as being made sequentially: first the farmer decides whether or not to produce opium, and then he decides on the amount. This way of thinking about the choices is somewhat counter to how I presented them earlier, but the important thing here is not the tree structure of the problem implied by the nested logit, it is to analyze the correlation structure among the random error terms.

The nested logit probabilities are found in the following way. Let the distribution of the error terms be given by the generalized extreme value distribution (the derivation is from Dagsvik (2006))

$$\Pr\left[\epsilon_0 \leq x_0, \ldots, \epsilon_J \leq x_J\right] = F(\boldsymbol{x}) = \exp\left(-G(\exp(-x_0), \exp(-x_1), \ldots, \exp(-x_J))\right)$$

where $G(\cdot)$ follows the standard assumption in McFadden (1978). $F(\boldsymbol{x})$ is a joint distribution function whose one–dimensional marginal distributions are extreme value distributions (McFadden, 2001). If the utility function is an additative RUM (as in equation (18) here), McFadden (1978) proved that

$$\Pr(U_{ij} = \max_k(U_{ik})) = \frac{\exp(V_{ij})G_j(\exp(V_{i1}), \ldots, \exp(V_{iJ}))}{G(\exp(V_{i1}), \ldots, \exp(V_{iJ}))}$$

If we let

$$G(x) = x_0 + \left(x_1^{1/\tau} + x_2^{1/\tau}\right)^{\tau}$$

we have the nested model with correlation between alternative 1 and 2, $\rho(\epsilon_1, \epsilon_2) = 1 - \tau^2$, and no correlation across nests, $\rho(\epsilon_0, \epsilon_j) = 0$ for $j = 1, 2$. Following McFadden (1978) and Dagsvik (2006) we find that the choice probabilities are given by

$$P_{i0} = \frac{\exp(V_{i0})}{\exp(V_{i0}) + [\exp(V_{i1})^{1/\tau} + \exp(V_{i2})^{1/\tau}]^{\tau}}$$

and

$$P_{ij} = \frac{\left[\exp(V_{i1})^{1/\tau} + \exp(V_{i2})^{1/\tau}\right]^{\tau-1}\exp(V_{ij}/\tau)}{\exp(V_{i0}) + (\exp(V_{i1})^{1/\tau} + \exp(V_{i2})^{1/\tau})^{\tau}}$$

for $j = 1, 2$. We immediately see that if $\tau = 1$ the standard multinomial case emerges. Hence the multinomial logit model is a special case of the nested model, and testing whether $\tau \neq 1$ will show whether using the non–nested multinomial model is appropriate or not.

Importantly, the Stata `nlogit` command, which I have used in the estimation, estimates Daly (1987)'s non–normalized nested logit model (NNNL), which departs from the McFadden (1974)'s random utility maximization nested logit model (RUMNL) in that the inclusive value terms are not normalized by the dissimilarity parameter(s) (Heiss, 2002). Here τ is the dissimilarity parameter. Koppelman and Wen (1998) shows that only in very special cases the NNNL model is consistent with random utility maximization, and they also shows that estimates of the coefficients may be very different under the two different models. Heiss (2002), however, shows that if only alternative–specific coefficients are estimated (as here), then the model is consistent with utility maximization. He also shows that in this case the log likelihood and the dissimilarity parameter(s) will be the same as if the model was estimated using the McFadden (1974) approach, but the coefficients will be scaled differently (they will be proportional to the coefficients from the RUMNL model by a factor equal to the inverse of the dissimilarity parameter). As we only include the nested multinomial logit model as a test on whether the normal non–nested approach used earlier is sufficient, and this test only uses the calculated dissimilarity parameter and the estimated log likelihood, which are similar in the two models given that only alternative–specific parameters are estimated, I do not go deeper into this.

The test of the assumed independence of the error terms across alternatives, is a likelihood ratio test of whether the dissimilarity parameter is significantly different from 1. The test statistics is

$$\mathrm{LR} = -2\left[\ln\left(L\left(\boldsymbol{\beta}^{\mathrm{R}}\right)\right) - \ln(L(\boldsymbol{\beta}))\right] \sim \chi^2_{\text{number of restrictions}} \tag{19}$$

where $\ln(L(\boldsymbol{\beta}^{\mathrm{R}})$ is the log likelihood for the model with restrictions and $\ln(L(\boldsymbol{\beta}))$ is the log likelihood for the unrestricted model. Here we impose one restriction, and if we choose significance level $\alpha = .05$, the critical value is 3.84. As we will see in the next section, we can not reject the null hypothesis that the dissimilarity parameter is equal to one in any of the specifications, hence the multinomial approach is sufficient.

The results from the nested logit models can be found in Table 8. The purpose of including this specification is to test whether the multinomial approach is valid, or whether the imposed IIA assumption violates the correlation structure in the random error terms.

To be able to estimate the model in Stata, I have had to convert my previous multinomial logistic regression to a conditional logit form, using alternative–specific dummies. For this reason I also include the previous multinomial logistic regression results in Table 8 (regressions (1) and (2)), as a test that the inclusion of the dummy variables does not alter any of the results (and as a test that my computer code does not include any mistakes). Column (3) and (4) of the table gives the nested logit estimates.

The likelihood ratio test gives the following results: for the model without regional dummies, the test statistic from (19) becomes

$$\hat{\text{LR}} = -2[-634.49 - -633.73] = 1.52 < 3.84$$

For the model with regional dummies,

$$\hat{\text{LR}} = -2[-603.01 - -602.99] = 0.04 < 3.84$$

Hence the test cannot reject the null hypotheses that $\tau = 1$. For this reason we safely continue to use the non–nested multinomial model rather than the nested.

Table 8: Nested logit estimates

		(1)			(2)			(3)			(4)		
Variables		Coef.		SE	Coef.		SE	Coef.		SE	Coef.		SE
const ×	Polycrop	-.51147		(.630)	-2.081076	*	(1.027)	-.179891		(.271)	-1.744335		(1.703)
	Monocrop	-3.08238	***	(.805)	-2.699812	*	(1.229)	-2.826832		(.548)	-2.357403		(1.855)
Own in jeribs ×	Polycrop	.025015	**	(.010)	.025252	*	(.011)	.014185	**	(.006)	.021869		(.017)
	Monocrop	-.105485	***	(.023)	-.107302	***	(.024)	-.114715		(.020)	-.110506	***	(.027)
Outstanding loan ×	Polycrop	.00449		(.006)	.002309		(.005)	.001447		(.003)	.001794		(.005)
	Monocrop	.007364		(.006)	.005015		(.005)	.004253		(.003)	.004474		(.005)
Loan 03, not family ×	Polycrop	.002679		(.002)	.002001		(.002)	.000748		(.001)	.001551		(.003)
	Monocrop	.009009	**	(.003)	.007404	*	(.003)	.006742	**	(.002)	.006928	*	(.003)
Price of vegetables ×	Polycrop	.201512	**	(.074)	.226104	***	(.070)	.076261		(.050)	.182833		(.199)
	Monocrop	.209862	*	(.083)	.243306	***	(.081)	.077909		(.066)	.198481		(.211)
Price of wheat ×	Polycrop	.074009		(.068)	.239755	**	(.087)	.013468		(.033)	.190417		(.228)
	Monocrop	.19622	**	(.073)	.362036	***	(.092)	.138189	***	(.043)	.31301		(.228)
Price of labor ×	Polycrop	-.002881	***	(.001)	-.002411	***	(.001)	-.00111		(.001)	-.001964		(.002)
	Monocrop	-.002347		(.002)	-.001449		(.002)	-.000365		(.002)	-.000934		(.003)
Price of opium ×	Polycrop	.000057	*	(.000)	.000051		(.000)	.000018		(.000)	.000043		(.000)
	Monocrop	.000095	***	(.000)	-.000007		(.000)	.000061	*	(.000)	-.000016		(.000)
Eradication in 2003 ×	Polycrop	.157325		(.297)	.509381		(.309)	-.013799		(.132)	.388		(.590)
	Monocrop	.757856	*	(.342)	1.0502	***	(.361)	.582847	**	(.217)	.929943		(.613)
Regional dummies		No			Yes			No			Yes		
τ		—			—			2.604634		(1.387)	1.232673		(1.263)
Log likelihood		-634.49			-603.01			-633.73			-602.99		
Iterations		6			7			306			12		
Obs		909			909			909			909		

* significant at 5%; ** significant at 1%; *** significant at .5%

6 Conclusions

This master thesis has analyzed whether or not opium production in Afghanistan is debt–induced, i.e. whether the cropping strategy of indebted farmers are different from other farmers. The main reason for studying this is that, to my knowledge, a thorough statistical analysis of this question has not yet been done. Many studies of a more qualitative nature have examined the question, and there seems to be a broad consensus in the literature that credit and opium production is interlinked (Mansfield, 1999, 2001, 2004a; Rubin, 2004; UNODC, 1999b, 2003b). However, this story is far from as obvious as it may seem to be by looking at correlation tables between debt and opium production, where the conditional probability of being an opium producer is found to significantly higher if the respondend has debt (see section 4). When controlling for price incentives, social class and eradication risk, we find no significant effect of debt on the conditional probability of being an opium producer (table 5). It is interesting to note that price incentives, especially prices on other agricultural crops, and eradication are important determinants of opium production. Eradication in the village in 2003 significantly *increases* the probability of opium production in 2004. Several possible explanations to this have been given in Section 5.1, among them is that eradication may free labor, which is important since opium production is extremely labor–intensive. We also find that the more land the farmer owns, i.e. the wealthier he is, the higher is the probability that he is an opium producer. We hence find a somewhat different picture than what is usually presented, here the opium farmer responds to price incentives, he is not in the poorest segments of the population, and debt seems to be unimportant for the choice of cropping strategy.

The reason why we do not find that the conditional probability depends on debt, is that farmers have two different (but often overlapping) incentives to produce opium. First, as it is the most valuable crop, farmers produce it to maximize the household's utility. But opium production is both illegal and anti–Islamic, hence there exist both social and moral costs to producing opium, which constrain the number of farmers choosing to produce opium for profit. Second, farmers produce opium to obtain loans, as opium production often is a prerequisite to be granted salaam loans, loans that are to be repaid in kind.

When we introduce moral costs into our theoretical model, we find that the heterogeneity in moral costs creates two subpopulations of opium farmers, the "opportunists" that maximize profit and the "moralists" that produces opium out of necessity. We also find that the two groups will have a different cropping strategy, which makes us able to distinguish them empirically: the "moralists" will produce zero opium until their accumulated debt reaches a tipping point, and when this point is reached they will produce at maximum capacity. The "opportunists" however, will find it optimal to diversify their production, producing opium as well as other crops.

To separate the two effects we therefore estimate choice probabilities for ending up as a monocropping opium farmer, relative to producing some opium or not producing opium. This model gives very intuitive results: for the monocropping opium farmers debt is an important determinant, while for the others debt is unimportant. We also find that the cost of labor is important for the “opportunists”, while not important for the “moralists”. Finally we find that eradication in the village is an important determinant of opium production only for the “moralists”, supporting that these farmers have experienced a bad shock that have increased their debt beyond the tipping point.

References

ACBAR (2000): *Helmand Initiative Socio-Economic Survey*Agency Coordinating Body for Afghan Relief Survey Unit.

Cameron, C. A., and P. K. Trivedi (2005): *Microeconometrics: Methods and Applications*. Cambridge University Press, New York.

Clemens, J. P. (2005): "Opium in Afghanistan: Prospects for the Success of Source-Country Drug Control Policies," Master's thesis, Harvard University.

Dagsvik, J. K. (2006): *Probabilistic Models for Qualitative Choice Behavior*. University of Oslo.

Daly, A. (1987): "Estimating "tree" logit models," *Transportation Research Part B: Methodological*, 21(4), 251–267.

Heiss, F. (2002): "Specification(s) of Nested Logit Models," Discussion paper, MEA - Mannheim Research Institute for the Economics of Aging.

Koppelman, F. S., and C.-H. Wen (1998): "Alternative nested logit models: structure, properties and estimation," *Transportation Research Part B: Methodological*, 32(5), 289–298.

Luce, D. R. (1959): *Individual Choice Behavior*. Wiley, New York.

Mansfield, D. (1999): "Alternative development: the modern thrust of supply-side policy," *Bulletin on Narcotics*, LI(1), 19–44.

——— (2001): "The Economic Superiority of Illicit Drug Production: Myth and Reality. Opium Poppy Cultivation in Afghanistan," in *International Conference on Alternative Development in Drug Control and Cooperation, Feldafing.*

——— (2004a): "Diversity and Dilemma: Understanding Rural Livelihoods and Addressing the Causes of Opium Poppy Cultivation in Nangarhar and Laghman, Eastern Afghanistan," Discussion paper, Project for Alternative Livelihoods (PAL) in Eastern Afghanistan.

——— (2004b): "What is Driving Opium Poppy Cultivation? Decision Making Amongst Opium Poppy Cultivators in Afghanistan in the 2003/4 Growing Season," in *Second Technical Conference on Drug Control Research*. UNODC/ONDCP.

——— (2005): "What is Driving Opium Poppy Cultivation? The Pressures to Reduce Opium Poppy Cultivation in Afghanistan in the 2004/05 Growing Season," Discussion paper, Afghan Drugs Inter Departmental Unit of the UK Government.

——— (2006): "Exploring the "Shades of Grey": An Assessment of the Factors Influencing Decisions to Cultivate Opium Poppy in 2005/06," *Afghan Drugs Inter Departmental Unit (UK Government).*

Mansfield, D., and A. Pain (2006): "Opium Poppy Eradication: How to raise risk when there is nothing to lose?," Discussion paper, Afghanistan Research and Evaluation Unit.

McFadden, D. (1974): "Conditional Logit Analysis of Qualitative Choice Behavior," in *Frontiers in econometrics*, ed. by P. Zarembka, pp. 105–142. Academic Press, New York.

——— (1978): "Modelling the Choice of Residential Location," in *Spatial Interaction Theory and Planning Models*, ed. by A. Karlqvist, L. Lundqvist, F. Snickars, and J. Weibull, pp. 75–96. North Holland, Amsterdam.

——— (2001): "Economic Choices," *The American Economic Review*, 91(3), 351–378.

Molla, D. (2003): "Food Aid, Wheat Prices and Poppy Cultivation in Afghanistan: Is There a Link?," Discussion paper, World Food Programme.

Rashid, A. (2002): *Taliban: Islam, Oil and the New Great Game in Central Asia.* I B Tauris & Co Ltd, London.

Rubin, B. R. (2004): "Road to Ruin: Afghanistan's Booming Opium Industry," Center on International Cooperation, New York University.

Train, K. E. (2003): *Discrete Choice Methods with Simulation.* Cambridge University Press.

UNODC (1999a): "Access to Labour: The Role of Opium in the Livelihood Strategies of Itinerant Harvesters Working in Helmand Province, Afghanistan," *Strategic Study Series*, 4.

UNODC (1999b): "The Role of Opium as a source of Informal Credit," *Strategic Study Series*, 3.

UNODC (2003a): *Afghanistan Opium Survey 2003* United Nations, Office on Drugs and Crime.

UNODC (2003b): *The opium economy in Afghanistan—An international problem* United Nations, Office on Drugs and Crime, New York.

UNODC (2004): *Farmers Intentions Survey 2003/2004* United Nations, Office on Drugs and Crime.

UNODC (2006): *World Drug Report*. United Nations Office on Drugs and Crime, Vienna, Austria.

A Computer code

I have used Stata for estimation. The code used to create Tables 5–8 is given below. The trick of using constraints to avoid specifing the upper level utility in the nested logit model is from Heiss (2002).

A.1 Binary logit

```
use "FIS_full.dta";

egen opiumprice1=mean(opiumprice), by(id_region);
replace opiumprice=opiumprice1 if opiumprice==.;

egen eradication1=mean(eradicationin2003), by(id_district);
replace eradicationin2003=eradication1 if eradicationin2003==.;
replace eradicationin2003=0 if eradicationin2003==.;
drop eradication1;

gen loan_notfam=loan_amount_2003-_29_amount_Family;
egen arbeidsdag1=mean(arbeidsdag);
replace arbeidsdag=arbeidsdag1 if arbeidsdag==.;

replace loan_notfam=loan_notfam/1000;
replace outstandingloan_amount=outstandingloan_amount/1000;
replace loan_amount_2003=loan_amount_2003/1000;

log using "simple_logit", replace text;
logit opium2004_dummy owninjeribs loan_amount_2003 outstandingloan_amount vegetableprice wheatprice arbeidsdag opiumprice eradicationin2003
if farmer_production_2004>0 & farmer_production_2004!=.;
logit opium2004_dummy owninjeribs loan_amount_2003 outstandingloan_amount vegetableprice wheatprice arbeidsdag opiumprice eradicationin2003
regionc1-regionc4 if farmer_production_2004>0 & farmer_production_2004!=.;
logit opium2004_dummy owninjeribs loan_notfam outstandingloan_amount vegetableprice wheatprice arbeidsdag opiumprice eradicationin2003
if farmer_production_2004>0 & farmer_production_2004!=.;
logit opium2004_dummy owninjeribs loan_notfam outstandingloan_amount vegetableprice wheatprice arbeidsdag opiumprice eradicationin2003
regionc1-regionc4 if farmer_production_2004>0 & farmer_production_2004!=.;

log c;
```

A.2 Tobit regression

```
use "FIS_full.dta";

egen opiumprice1=mean(opiumprice), by(id_region);
replace opiumprice=opiumprice1 if opiumprice==.;

egen eradication1=mean(eradicationin2003), by(id_district);
replace eradicationin2003=eradication1 if eradicationin2003==.;
replace eradicationin2003=0 if eradicationin2003==.;
drop eradication1;

gen loan_notfam=loan_amount_2003-_29_amount_Family;
egen arbeidsdag1=mean(arbeidsdag);
replace arbeidsdag=arbeidsdag1 if arbeidsdag==.;

gen rel04=poppy_production_2004/farmer_production_2004;
gen poppycat=.;
replace poppycat=0 if farmer_production_2004>0 & farmer_production_2004!=. & poppy_production_2004==0;
replace poppycat=1 if farmer_production_2004>0 & farmer_production_2004!=. & poppy_production_2004>0 & rel04<1;
replace poppycat=2 if farmer_production_2004>0 & farmer_production_2004!=. & poppy_production_2004>0 & rel04==1;
replace loan_amount_2003=loan_amount_2003/1000;
replace loan_notfam=loan_notfam/1000;
replace outstandingloan_amount=outstandingloan_amount/1000;

log using "tobit", replace text;
tobit poppy_production_2004 owninjeribs outstandingloan_amount loan_amount_2003 vegetableprice wheatprice arbeidsdag opiumprice eradicationin2003
if farmer_production_2004>0 & farmer_production_2004!=., ll;
outreg using tobit, se replace;

tobit poppy_production_2004 owninjeribs outstandingloan_amount loan_amount_2003 vegetableprice wheatprice arbeidsdag opiumprice eradicationin2003
regionc1-regionc4 if farmer_production_2004>0 & farmer_production_2004!=., ll;
outreg using tobit, se append;
```

```
tobit poppy_production_2004 owninjeribs outstandingloan_amount loan_notfam vegetableprice wheatprice arbeidsdag opiumprice eradicationin2003
if farmer_production_2004>0 & farmer_production_2004!=., ll;
outreg using tobit, se append;

tobit poppy_production_2004 owninjeribs outstandingloan_amount loan_notfam vegetableprice wheatprice arbeidsdag opiumprice eradicationin2003
regionc1-regionc4 if farmer_production_2004>0 & farmer_production_2004!=., ll;
outreg using tobit, se append;

log c;
```

A.3 Multinomial logit

```
use "FIS_full.dta";

egen opiumprice1=mean(opiumprice), by(id_region);
replace opiumprice=opiumprice1 if opiumprice==.;

egen eradication1=mean(eradicationin2003), by(id_district);
replace eradicationin2003=eradication1 if eradicationin2003==.;
replace eradicationin2003=0 if eradicationin2003==.;
drop eradication1;

gen loan_notfam=loan_amount_2003-_29_amount_Family;
egen arbeidsdag1=mean(arbeidsdag);
replace arbeidsdag=arbeidsdag1 if arbeidsdag==.;

gen rel04=poppy_production_2004/farmer_production_2004;
gen poppycat=.;
replace poppycat=0 if farmer_production_2004>0 & farmer_production_2004!=. & poppy_production_2004==0;
replace poppycat=1 if farmer_production_2004>0 & farmer_production_2004!=. & poppy_production_2004>0 & rel04<1;
replace poppycat=2 if farmer_production_2004>0 & farmer_production_2004!=. & poppy_production_2004>0 & rel04==1;
replace loan_amount_2003=loan_amount_2003/1000;
replace loan_notfam=loan_notfam/1000;
replace outstandingloan_amount=outstandingloan_amount/1000;

log using "multinomial_logit", replace text;
mlogit poppycat owninjeribs outstandingloan_amount loan_notfam vegetableprice wheatprice arbeidsdag opiumprice eradicationin2003
regionc1-regionc4, basecategory(0);
mlogit poppycat owninjeribs outstandingloan_amount loan_notfam vegetableprice wheatprice arbeidsdag opiumprice eradicationin2003, basecategory(0);
mlogit poppycat owninjeribs outstandingloan_amount loan_amount_2003 vegetableprice wheatprice arbeidsdag opiumprice eradicationin2003
regionc1-regionc4, basecategory(0);
mlogit poppycat owninjeribs outstandingloan_amount loan_amount_2003 vegetableprice wheatprice arbeidsdag opiumprice eradicationin2003, basecategory(0);
log c;
```

A.3.1 Nested multinomial logit

```
use "FIS_full.dta";

egen opiumprice1=mean(opiumprice), by(id_region);
replace opiumprice=opiumprice1 if opiumprice==.;

egen eradication1=mean(eradicationin2003), by(id_district);
replace eradicationin2003=eradication1 if eradicationin2003==.;
replace eradicationin2003=0 if eradicationin2003==.;

gen loan_notfam=loan_amount_2003-_29_amount_Family;
egen arbeidsdag1=mean(arbeidsdag);
replace arbeidsdag=arbeidsdag1 if arbeidsdag==.;

drop opiumprice1 eradication1 arbeidsdag1 provc1-provc29 id_district
id_province id_region id_village opiumprice_mean_village opiumprice_mean_district
opiumprice_difference_village opiumprice_difference_district outstandingloanyear;

gen rel04=poppy_production_2004/farmer_production_2004;
gen poppycat=.;
replace poppycat=0 if farmer_production_2004>0 & farmer_production_2004!=. & poppy_production_2004==0;
replace poppycat=1 if farmer_production_2004>0 & farmer_production_2004!=. & poppy_production_2004>0 & rel04<1;
replace poppycat=2 if farmer_production_2004>0 & farmer_production_2004!=. & poppy_production_2004>0 & rel04==1;
replace loan_notfam=loan_notfam/1000;
replace outstandingloan_amount=outstandingloan_amount/1000;

expand 3;
sort id;
gen seq=mod(_n-1,3);
gen valg=0;
replace valg=1 if poppycat==0 & seq==0;
```

```
replace valg=1 if poppycat==1 & seq==1;
replace valg=1 if poppycat==2 & seq==2;

gen asc_1=0;
gen asc_2=0;
replace asc_1=1 if seq==1;
replace asc_2=1 if seq==2;

gen owninjeribs_1=owninjeribs*asc_1;
gen owninjeribs_2=owninjeribs*asc_2;

gen outstandingloan_amount_1=outstandingloan_amount*asc_1;
gen outstandingloan_amount_2=outstandingloan_amount*asc_2;

gen loan_notfam_1=loan_notfam*asc_1;
gen loan_notfam_2=loan_notfam*asc_2;

gen vegetableprice_1=vegetableprice*asc_1;
gen vegetableprice_2=vegetableprice*asc_2;

gen wheatprice_1=wheatprice*asc_1;
gen wheatprice_2=wheatprice*asc_2;

gen arbeidsdag_1=arbeidsdag*asc_1;
gen arbeidsdag_2=arbeidsdag*asc_2;

gen opiumprice_1=opiumprice*asc_1;
gen opiumprice_2=opiumprice*asc_2;

gen eradicationin2003_1=eradicationin2003*asc_1;
gen eradicationin2003_2=eradicationin2003*asc_2;

gen regionc1_1=regionc1*asc_1;
gen regionc1_2=regionc1*asc_2;

gen regionc2_1=regionc2*asc_1;
gen regionc2_2=regionc2*asc_2;

gen regionc3_1=regionc3*asc_1;
gen regionc3_2=regionc3*asc_2;

gen regionc4_1=regionc4*asc_1;
gen regionc4_2=regionc4*asc_2;

egen nothing=fill(1 0 0 1 0 0);
cons define 1 nothing = 0;

nlogitgen o_vs_not_o = seq(notop:0,op:1|2);

log using "nested_logit", replace text;

clogit valg asc_* owninjeribs_* outstandingloan_amount_* loan_notfam_* vegetableprice_* wheatprice_* arbeidsdag_* opiumprice_* eradicationin2003_*
regionc1_* regionc2_* regionc3_* regionc4_*, group(id);
nlogit valg (seq= asc_* owninjeribs_* outstandingloan_amount_* loan_notfam_* vegetableprice_* wheatprice_* arbeidsdag_* opiumprice_* eradicationin2003_*
regionc1_* regionc2_* regionc3_* regionc4_*)(o_vs_not_o=nothing), group(id) const(1) ivc(notop=1 op=1);

clogit valg asc_* owninjeribs_* outstandingloan_amount_* loan_notfam_* vegetableprice_* wheatprice_* arbeidsdag_* opiumprice_* eradicationin2003_*, group(id);
nlogit valg (seq= asc_* owninjeribs_* outstandingloan_amount_* loan_notfam_* vegetableprice_* wheatprice_* arbeidsdag_* opiumprice_* eradicationin2003_*)
(o_vs_not_o=nothing), group(id) const(1) ivc(notop=1 op=1);

nlogit valg (seq= asc_* owninjeribs_* outstandingloan_amount_* loan_notfam_* vegetableprice_* wheatprice_* arbeidsdag_* opiumprice_* eradicationin2003_*
regionc1_* regionc2_* regionc3_* regionc4_*)(o_vs_not_o=nothing), group(id) const(1) ivc(notop=1);
nlogit valg (seq= asc_* owninjeribs_* outstandingloan_amount_* loan_notfam_* vegetableprice_* wheatprice_* arbeidsdag_* opiumprice_* eradicationin2003_*)
(o_vs_not_o=nothing), group(id) const(1) ivc(notop=1);

log c;
```