DISCUSSION PAPERS
Department of Economics
University of Copenhagen

05-27

Commercialisation and Poverty in Tanzania: Household-level Analysis

Elina Eskola

Studiestræde 6, DK-1455 Copenhagen K., Denmark
Tel. +45 35 32 30 82 - Fax +45 35 32 30 00
http://www.econ.ku.dk

Commercialisation and Poverty in Tanzania:

Household-level Analysis

ELINA ESKOLA[12]

Abstract:

Data from a recent Tanzanian household survey are used to investigate households' connectedness to market economy i.e. commercialisation. The study puts emphasis on facilitating access to the nearest markets and market information as means to enhance commercialisation. Commercialisation and total consumption are found to be highly correlated and mutually reinforcing. The more commercialised the household is, the higher probability it has to be well off and thus enhanced commercialisation should be encouraged as a way to increase households' welfare. The distance to the nearest market and the availability of market information are found to be significant factors in households' degree of commercialisation. However, the importance of large regional fixed effects highlights the importance of the structural barriers for trade in Tanzania.

Keywords:
Commercialisation, poverty, transaction cost, household model, Tanzania.

JEL classification: C31, I31, Q12

[1] University of Sussex, currently visiting the University of Copenhagen. Address for correspondence: Elina.Eskola@econ.ku.dk, Studiestraede 6, DK-1455 Copenhagen, Denmark.
[2] I gratefully acknowledge the useful comments from Alan Winters, Andrew Newell and Barry Reilly in writing this paper.

1 Introduction

The underlying assumption of economic theory states that trade increases people's welfare by allowing them to specialise and capture the producer and consumer surplus. Nevertheless, large numbers of households in the developing world still live largely in subsistence farming and the puzzle of rational decision makers withdrawing from the markets has been part of the development debate for years (see e.g. McKay *et al.* 1997). This withdrawal, however, is rarely done by choice. People resort to subsistence farming when the market institution fails to function efficiently, which leads to suboptimal specialisation and lower levels of welfare as the gains from trade are foregone. The market institution fails to facilitate the trade when transaction costs due to market exchange create disutility greater than the utility gained through the market transaction so that no market transaction takes place. Rather than a complete failure of the market, a more general case is that market transactions are replaced by surrogate institutions, such as informal reciprocal crop sharing, or that markets in fact exist but only the most efficient households are able to use them. Poor infrastructure, inefficient marketing systems, lack of information and the risk involved in trading are factors that increase the cost of trade. The higher the transaction costs, the higher the benefits from trade must be before a household is willing to engage in trade.

De Janvry *et al*. (1991) argue that substantial benefits would be gained if missing markets could be (re-)opened for trade by decreasing the costs of transaction. The recommendation for reducing the obstacles for private households to participate especially in food market e.g. via reduced trade restrictions and better roads is well pronounced also elsewhere in the literature (Fafchamps 1997). In recent years, the argument in favour of facilitating market access for the purpose of poverty alleviation has regained sense of urgency within the donor community as well as within the developing country governments. For example the Tanzanian government has set efficient agricultural marketing high on the agenda of improving welfare in the country (Tanzanian PRSP, Agricultural Marketing Policy), and a private sector development programme (BEST) has already been launched in order to achieve this goal.

Even though the general trend in the development discourse is in favour of commercialisation, how well it works in different contexts depends on the characteristics of the households, potential of the local market as well as the legal and political framework in which the commercialisation is to take

place. Furthermore, the success of further commercialisation in relieving poverty depends on how well the households are currently integrated in the market and to what extent the opportunities provided by specialisation have already been exploited. This study aims at providing evidence of the magnitude and nature of the welfare impact of commercialisation in Tanzania. Special attention is paid to the role of market access and market information as a constraint to commercialisation, and their role and potential in poverty alleviation. The purpose of the chapter is to analyse the current situation in the country and to establish and measure the links between commercialisation and consumption at the household level. The central hypothesis of this study is that access to markets and market information is a significant factor in a household's decision on the degree to which the household participates in trade. Increased participation in trade at the local market is likely to increase consumer/producer surplus for the participating household and thus increase the household's welfare. Hence, better market access is likely to increase the consumption level of the household thorough increased levels of trade. The study is aiming to answer two central research questions:

- How does commercialisation affect households' wellbeing?
  and
- What are the determinants of commercialisation of the household economy in Tanzania?

The remaining chapter is organised as follows: part two reviews the theoretical debate on commercialisation and its links to poverty, as well as the welfare impact of better market access; part three presents the theoretical model employed, and the empirical approach; part four introduces the data used for the empirical analysis, as well as discusses the construction of the main variables; part five explains the econometric methodology; part six presents and discusses the empirical results; and finally part seven concludes.

2 Conceptual Framework of Commercialisation

The term most often used on gains from prior subsistence farms engaging into trade is 'commercialisation' that implies increased market transactions for capturing the gains from specialisation (von Braun 1995). By definition, commercialisation can occur both on the output side of production with increased marketed surplus or on the input side with increased use of purchased inputs, and most often these commercialisation effects occur simultaneously. Not only producers

but also consumers benefit from trading through consumer surplus and further ability to specialise in their own production. The term commercialisation can also be used to refer to market integration of household economy so that larger part of consumption is acquired through market transactions, which usually leads to further specialisation in the use of household's productive resources (Von Braun & Kennedy 1994, 11-12). A broad definition of commercialisation referring to market integration by engaging in trade is adopted in this study.

2.1 Commercialisation and Poverty

The links between trade and poverty have been widely explored in the literature. A substantial contribution to the debate has been the survey edited by von Braun and Kennedy (1994) that summarises the results from several studies about commercialisation in 11 different sites in the developing world. In addition to reviewing case studies the authors lay out a framework for the possible ways in which commercialisation and welfare, in particular income and nutrition, are linked and could be analysed (figure 1).

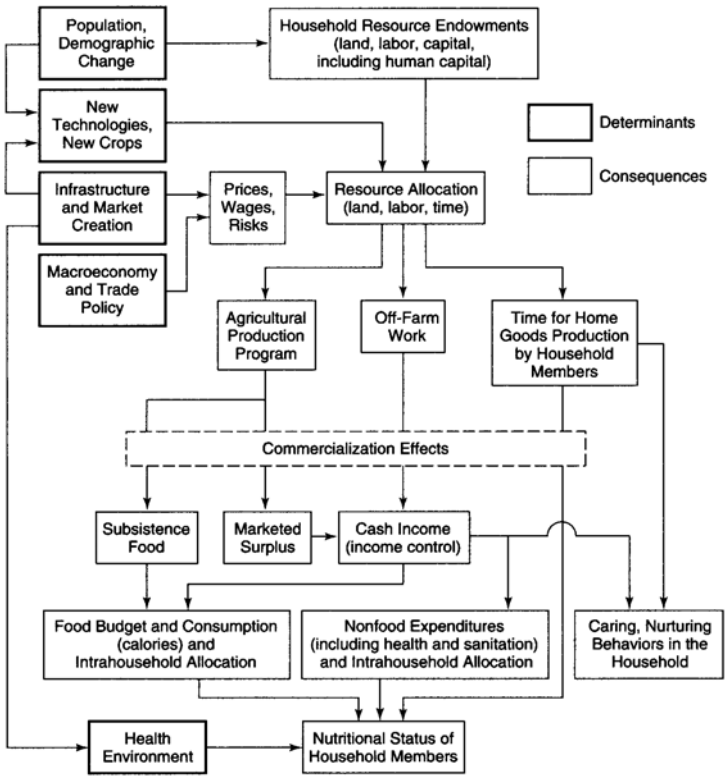

Figure 1: Commercialization at the household level: determinants and consequences for income and nutrition. *Source*: von Braun and Kennedy (1994) p. 13.

Among the most important *exogenous determinants* of commercialisation are population change, availability of new technologies, infrastructure and market creation, and macroeconomic and trade policy. Some of these factors have more immediate effect on the farmer's decision while others have more long-term effects. At the household level, the key factors in the commercialisation process are the availability of improved seeds and agricultural practices as well as investment in infrastructure and policies for market creation. All of the determinants of commercialisation are intertwined and mutually reinforcing, and thus improved trade policies, technical change, and better infrastructure are often inseparable tools and/or consequences of a commercialisation intervention. How the possible increased welfare of the household finally translates into increased welfare of the individuals depends on the decision making process within the household (*endogenous consequences* of commercialisation).

Making generalising conclusions about the welfare impact of a commercialization programme is difficult. However, von Braun and Kennedy argue that in general overcoming the market failure that leads to subsistence farming is likely to cause beneficial outcomes through several links. The reduction of transaction cost creates a stronger price incentive for a producer to engage in trade. Improved markets and transportation networks increase the number of suppliers, which is also likely to lead to more reliable supply of food crops and less volatile prices at the markets. This, on the other hand, would lower the risk involved in trade and allow otherwise risk-avert and vulnerable households to specialise and benefit from selling cash crops as well as consuming larger variety of goods acquired from the market. It is important to bear in mind, however, that commercialization does not only relate to selling cash crops but commercialisation of food crops is an ever more important part of enhanced livelihoods in poor countries (see e.g. Heltberg & Tarp 2002). Furthermore, cash cropping does not need to be competitive with other agricultural activities but food cropping and cash cropping often grow and decline together (Maxwell 2001). Usually production for the market is only done when the household basic demand for food has been ensured. Poor households are willing to deviate from specialising in profit maximising resource allocation to maintain household food security based on own production of food crops when insurance markets are absent or unable to cover negative shocks from own crop failure or smooth consumption during a negative price shock at the market (von Braun & Kennedy 1994, 52). However, investment in improved technology can help subsistence farmers commercialize in low-risk ways. In order to facilitate their use and to allow for full benefits from commercialization, the

development of the financial and insurance markets is a crucial complement to the commercialization efforts. The smallest households who currently participate less than proportionally in the market are likely to benefit most from targeted commercialization efforts (von Braun & Kennedy 1994, 370).

Commercialisation of agriculture especially combined with expanded processing and trading activities has been observed to lead to a substantial expansion of demand for hired labour, which contributes to the income earning possibilities of poor households. To the extent that hired labour households rank among the malnourished poor, this employment effect may well be of particular benefit as demonstrated in case studies from Bangladesh, Indonesia, Guatemala, and Papua New Guinea (von Braun and Kennedy 1994, chapters 8, 11, 12 and 14). However, in case the new technology introduced to cultivate cash drops for the market is less labour intensive than the previous method and processing of the produce is not done in the village the effect on the demand for labour might also be negative. Still, selling new products and increased production for the market is likely to increase the household's gross incomes. However, the increase in the net income is likely to be less because of large substitution effects within agricultural production and between agriculture and off-farm employment. Often labour intensive cash crop production draws labour back into agriculture, and slows down the urbanization. Increased income leads most often to increased spending on food and other necessities by the poor. How the increased consumption and control over increased resources is distributed within the household is part of the intra-household decision-making process.

Contradictory to the framework outlined by von Braun and Kennedy (1994), concern has been raised about the possible damaging effects of commercialisation. It has been argued that even though commercialisation could be beneficial for the economy as a whole, its impact on the poor households would be mainly negative (see e.g. Lambert 1982 contradicted by Heywood & Hide 1994). The empirical literature has produced both negative and positive conclusions even on the same commercialisation programmes and the evidence put forward has been at best inconclusive. Von Braun and Kennedy (1986) reviewed some of the available case studies and concluded that many of the studies showing negative effects on farming households have been conceptually flawed and based on false assumptions made in the studies. Despite the possible shortcomings of individual studies, the worry for the possible adverse effect on poor farmers has been influential in policy

design and a matter of debate even though empirical studies at large have failed to confirm this negative link. Longhurst (1988) reviewed empirical evidence and concluded that the results show no clear outcome either against or for commercialisation, but that the welfare effect of commercialisation depends on the way in which the programme is designed to suit the given context. He argues that promotion of non-food crops with long maturation periods and lumpy income flows from a narrow and unstable market might lead to adverse effects on households' welfare whereas enhancing the commercialised production and selling of food crops that are complementary to the existing farming system and that give steady flows of income from well established local markets is likely to increase the consumption and welfare of the households.

Even though judging all non-food products as bad commercialisation means over-simplifying the reality and might lead to abandoning otherwise viable policy options, Longhurst's characterisation raises an important point of poverty impacts into the commercialisation debate: the way in which commercialisation is implemented is likely to change the multifaceted welfare outcomes of further market integration, and a smooth transition from subsistence farming into market economy cannot be assumed. Significant equity and environmental consequences may be caused by the change in production and consumption patterns unless appropriate policies are in place to facilitate the transition (Pingali & Rosengrant 1995). Markets work always in interaction with other institutions and policies and governments and policy-makers have an important role to play in ensuring long run efficient use of resources without adversely affecting the poor in the short run.

2.2 Welfare Analysis on Facilitating Market Access

The current study places emphasis on investigating the role of market access and information in commercialisation and poverty alleviation. Rural households, especially the poor, often say that one reason they cannot improve their living standards is that they face difficulties with market access due to distance to markets and lack of roads (IFAD 2001, p.161). The literature supports the hypothesis that inadequate market access in absence of roads, telephones and other infrastructure decreases welfare and hinders trade (Bougheas et al. 1999). Several empirical studies have concluded that sufficient provision of infrastructure services is an important component for poverty alleviation as such (e.g. van de Walle 2002a) and a necessary prerequisite for efficient trading (e.g. Ndumbaro 1995, Larson & Deininger 2001). Consequently many infrastructure projects have been

justified by referring them to be working for poverty reduction directly and indirectly through local trade creation.

The link from infrastructure to improved standard of living works through several channels. Development of transport and communication infrastructure enhances people's mobility, and increases interaction and informal learning. The change is reflected in increased use of imported goods, rising income, and entrepreneurial development that increases the capacity to perceive and the ability to seize comparative advantages. Rural credit markets often offer credit in kind but infrastructural development can reduce the extent of in-kind credits by transforming them into monetary exchange at lower transaction costs, which facilitates credit market expansion. Also labour markets are affected due to improved mobility of labour and the market becomes less fragmented providing more opportunities for the household members. New opportunities become available also for traditional farmers as new ideas spread through increased interaction, marketing of new inputs becomes logistically cheaper, and both factor and product markets operate more efficiently in infrastructurally developed areas facilitating the adoption of more efficient production. More concretely infrastructural development leads to reduction of marketing costs, expands markets, and improves market operation. Improved transportation infrastructure allows production of perishable and transport-intensive products to expand. Better access also converts latent demand into effective commercial demand and further expands the markets. Often this leads to further specialisation and possibilities to exploit economies of scale. (Ahmed 1994)

The theory clearly predicts that infrastructure plays a critical role in commercialisation and welfare. The methods for actually verifying and quantifying the improvements are still developing to better capture the short and long term direct and indirect welfare affects of road building (Ravallion 2001, Gootaert 2002, van de Walle 2002a). Gibson and Rozelle (2002) investigated how effective access to infrastructure is in reducing poverty in the context of Papua New Guinea. The paper argues that as the rural poor have least access to infrastructure, they would be in the best position to gain from additional investment. Van de Walle (2002b) presents further evidence from Vietnam and argues that the benefits of roads are indirect and dependent on interactions with other investments, existing social and physical infrastructure, geography, community and household characteristics. Often road networks also have economy-wide effects: van de Walle found significant welfare effects of road construction at both household and commune level and was able to conclude that the strongest

impacts were among the poorest households. However, as infrastructure is an expensive development policy to pursue, more empirical evidence is needed to ensure that the impact of investment is not only positive but also superior to alternative uses of the resources. Fan et al. (2003) summarise the results from several individual studies conducted at the International Food Policy Research Institute (IFPRI) comparing different types of public investments to each other. They conclude that R&D, education, and infrastructure proved to be the most effective public spending in relation to poverty reduction. Also Datt and Joliffe (1999) produce empirical evidence of different welfare projects by analysing the effect of different determinants of poverty as well as their interaction with each other in Egypt. The authors argue that education and infrastructure are important complements and thus the benefits received from one depend on the level of achievement in the other.

Also evidence from Tanzania argues in favour of facilitating market access in order to increase welfare. Lucas et al. (1995) evaluated a regional road project between Njombe and Makete in Tanzania that included improvements of feeder roads, bridge construction and road maintenance. The impact study found a 70% increase in daily traffic and a heavy decrease in fare prices. They also observed increased participation in the markets and expansion in the geographical size of the market. Furthermore, attendance at health care facilities and political meetings increased with the easier access. Airey et al. (1989) on the other hand, evaluated the Songea-Makambo road in Tanzania. Also here the use of the road doubled or trebled and the cost of vehicle transportation halved after the construction of the road. Even though the school enrolment did not increase, it became easier to recruit teachers to schools, and also access to health care facilities improved significantly. Ndumbaro (1995) takes the analysis further down to the grass root level of the transportation problems in rural areas in Tanzania. The article was based on rapid rural appraisal (RRA) surveys in Songea Rural District. The author discussed particular problems that the farmers faced that are linked to transportation infrastructure but not directly solved by improved infrastructure. For example the farmers had problems with hiring trucks even when the road was built as the transportation markets were not yet liberalised at the time when the study was carried out. Ndumbaro also identified new possibilities related to but not directly caused by better roads, such as that access to markets not only boosted agricultural activities but also increased other products to be made and sold at the markets leading to diversification and decrease in the vulnerability of the farmers.

It is important to bear in mind that the benefits from improved infrastructure are conditional on appropriate policies being in place allowing and encouraging trade to develop. As Fafchamps et al. (2003) point out, the price transmission, i.e. how prices at the central markets are linked to the prices at more remote markets, is not only dependent on transport costs but also frictions in the marketing chain. Besides strict exchange rate policy at macro level and limited access to credit at micro level, natural barriers caused by the remote location of the producers, unreliability of the transportation services, and corruption are often important barriers for local trade. Even though the reasons for low market participation of small holders vary, the recommendation for reducing the obstacles for private households to participate, especially in food markets, remain unchanged in the literature. "Food market integration via reduced trade restrictions, better roads and transportation, and/or government food shops can be a powerful tool to boost cash crop production and to increase responsiveness of small farmers to price incentives" (Fafchamps 1997).

3 The model of Trade and Income

As discussed above, the literature calls for active measures of supporting commercialization of subsistence farming as the markets alone in remote areas cannot be expected to lead to the optimum level of exchange, specialisation and allocation of resources. Here a model is constructed to theoretically establish the link from transaction costs to poor market integration. The model is then empirically tested in order to quantify the impact of commercialisation on consumption, as well as to identify ways to enhance commercialisation by lowering the cost of engaging in trade.

The main emphasis of the more recent trend in the literature has been so-called non-separable household models where production and consumption decisions cannot be separated from each other but the household optimal production decisions depend on consumption. The most commonly cited sources of non-separability are transaction costs, thin markets, and risk aversion. One of the grounding studies on transaction costs in relation to household models was the model built by de Janvry *et al.* (1991) who discuss the phenomenon of missing markets. They argue that the definition of a market failure is household rather than commodity specific: in case the household is close enough to the market, have sufficient profit margin for its products, and in general is able to cover the transaction costs by the benefits from trading, it is likely to be better off when engaging in trade than otherwise. However, households in remote locations whose cost to access the market is high are often unable to cover the transaction costs to participate in trade even though in absence of the

cost of trade engaging in the market would be beneficial. This leads to non-existing or thin markets. The higher the transaction costs, the higher the benefits from trade must be before a household is willing to engage in trade. Another example of a transaction cost based models is the framework presented by Key et al. (2000). They argue that farmers face both fixed transaction costs that are invariant to the quantity of goods traded, and proportional transaction costs that lower the price effectively received by sellers. Thus the decision to engage in trading is done in two parts: first the household decides whether or not to participate considering the fixed transaction costs, and second if they decide to participate, how much would they sell. The framework of fixed and proportional transaction costs has been tested by Heltberg and Tarp (2002) who conclude that even though decreasing both proportional and fixed transaction costs are both important factors of the households' decision, facilitating the market access of the currently isolated households is likely to be a pro-poor policy.

The model concentrating on the existence of transaction costs as referred to in this study is formulated by Omamo (1998a, 1998b). He focuses on the proportional transaction costs but widens the argument from producers to cover also the net buyers. He argues that transaction costs form a wedge between buying and selling price, i.e. they raise the price for the buyer and decrease the price for the seller. Omamo (1998a) investigates links between transport costs and production patterns by adjusting the agricultural household model to incorporate costly trade. He uses the model to explain the production patterns in Kenya, that have previously been considered irrational, and argues that producing low-yielding food crops over high-yielding cash crops can be seen as optimal food import substitution in presence of high transportation costs. In his other paper (1998b) Omamo elaborates the discussion, and argues that choosing low-yielding food crops can be explained even in absence of risk or lack of technical feasibility of the high-yielding marketed varieties if the new technologies involve greater specialisation that expands income but raises transaction costs by more. He tests the model with a simplified numerical model and simulates household optimal behaviour as the distance to the market (used as a proxy for transportation costs) increases and finds that even reasonably short distances make a difference in the optimal strategy. Omamo concludes that in his case study "for a household residing more than four kilometres from a market centre, expenditure savings from diversification outweight income losses" (Omamo 1998b).

3.2 The Current Model

The model applied in this study is based on the basic model outlined by Singh *et al*. (1986) and Bardhan & Udry (1999). It has been modified along the lines of Omamo (1998a, 1998b) to incorporate an explicit analysis of transportation margins into the welfare analysis, making the model non-separable. The household is assumed to maximise its utility function:

$$U = u(C_F, C_{NF}, l) \qquad (1)$$

where

C_F is consumption of food which can be produced at home or bought from the market,

C_{NF} is consumption of non-food items that cannot be produced at home but have to be bought from the market (say soap, kettle, shoes), and

l is leisure.

The utility function is assumed to be well behaved: quasi-concave with positive partial derivatives. It is maximised subject to constraints, namely time constraint (2), production constraint (3), and income constraint (4). The total time available to the household must be equal to the time it spends on work and the time allocated for leisure:

$$T = H + l \qquad (2)$$

where

T is the total time endowment and

H is time spent on work.

The production constraint, on the other hand, depicts the relation between inputs and outputs:

$$Q = f(L, A) \qquad (3)$$

where

Q is the total output of food that the household produces,

L is the total amount of labour (either family labour input or rented labour) used in production, and

A is the fixed endowment of other factors of production, such as land.

Furthermore, the income constraint is specified as follows:

$$\underbrace{\left(P_F + \lambda_1 - \lambda_2\right)\cdot\left(Q - C_F\right)}_{\text{Marketed surplus}} + \underbrace{W\left(H - L\right)}_{\text{Marketed labour}} = \left(P_{NF} + \lambda_1\right)C_{NF} \qquad (4)$$

where

P_F is the price of the food at the market,

P_{NF} is the price of the non-food item at the market,

λ_1 and λ_2 are the transaction costs of trading, such as going to the market, acquiring market information, and negotiating prices. For the simplicity of the model, this cost is assumed to be the same for all goods. The transaction cost λ_1 is added to the market price the household has to pay if the household is buying a good, and transaction cost λ_2 is deducted from the price that the household receives if the household is a net seller of food. For each household only one of the λ:s is relevant in terms of food depending on whether the household is a net buyer or a net seller. Unlike generally assumed in the household models, the cost of transaction does not have to be the same for the buyers and the sellers, but the incidence of the cost depends on the supply and demand elasticities. The status of the household is determined by the second term where the food consumption is deducted from the household's total food production. Finally,

W is the wage rate determining the value of the marketed labour. The wage is assumed to be a going market wage rate for which the household can either buy or sell labour. The whole term is negative if household uses more labour in its production that it gives, i.e. the household is net buyer of labour, and positive if household's own labour is higher than its labour use for production. If the household is neither buying nor selling labour the term for the marketed labour is zero. As the emphasis of the model is on trading goods, the possible transaction cost related to the labour market is ignored in the model.

All prices in the model, i.e. P_F, P_{NF}, and W are assumed to be exogenous and thus the household is a price taker. The constraints discussed above can also be combined into one budget constraint as follows:

$$\underbrace{\left(P_F+\lambda_1-\lambda_2\right)f\left(L,A\right)-W\cdot L+W\cdot T}=\underbrace{\left(P_F+\lambda_1-\lambda_2\right)\cdot C_F+\left(P_{NF}+\lambda_1\right)C_{NF}+W\cdot l} \quad (5)$$

Farm profit (Π)	Value of time	Value of total expenditure

The Lagrangian function for constrained optimisation then becomes:

$$L=u(C_F,C_{NF},l)+ \quad (6)$$
$$\mu\left[\left(P_F+\lambda_1-\lambda_2\right)f\left(L,A\right)-W\cdot L+W\cdot T-\left(P_F+\lambda_1-\lambda_2\right)\cdot C_F-\left(P_{NF}+\lambda_1\right)C_{NF}-W\cdot l\right]$$

The first order conditions for a net seller are:

$$\frac{\partial U}{\partial C_F}=\mu\left(P_F-\lambda_2\right)$$
$$\frac{\partial U}{\partial C_{NF}}=\mu\left(P_{NF}+\lambda_1\right) \quad (7)$$
$$\frac{\partial U}{\partial l}=\mu W$$
$$\left(P_F-\lambda_2\right)\frac{\partial Q}{\partial L}=W$$

and for a net buyer:

$$\frac{\partial U}{\partial C_F}=\mu\left(P_F+\lambda_1\right)$$
$$\frac{\partial U}{\partial C_{NF}}=\mu\left(P_{NF}+\lambda_1\right) \quad (8)$$
$$\frac{\partial U}{\partial l}=\mu W$$
$$\left(P_F+\lambda_1\right)\frac{\partial Q}{\partial L}=W$$

Some comparative statistics reveal how increase in the trading costs will change the household's behaviour. At optimum $C_F = C_F\left(P_F, P_{NF}, \lambda_1, \lambda_2, W, Y^*\right)$ and $C_{NF} = C_{NF}\left(P_F, P_{NF}, \lambda_1, \lambda_2, W, Y^*\right)$ where the income *Y** is defined as in the budget constraint:

$$Y^* = \left(P_F + \lambda_1 - \lambda_2\right)Q^* - W \cdot L^* + W \cdot T^* = \pi^* + W \cdot T^* \tag{9}$$

For the *net seller* of food crops, when trading costs go up, the household is likely to consume more of the food it produces at home and production is likely to decrease, i.e. the household will withdraw from the market. In order to see this more formally, differentiating C_F with respect to the trading cost λ_2 using the Slutsky equation will lead to:

$$\frac{\partial C_F}{\partial \lambda_2} = \frac{\partial C_F}{\partial \lambda_2}\Big|_{\Delta\pi^*=0} + \underbrace{\frac{\partial C_F}{\partial Y^*} \cdot \frac{\partial Y^*}{\partial \lambda_2}}_{\text{Profit effect}} = \underbrace{\frac{\partial C_F}{\partial \lambda_2}\Big|_{\Delta U=0}}_{>0} + \underbrace{\left(Q - C_F\right)}_{>0} \cdot \underbrace{\frac{\partial C_F}{\partial Y^*}}_{>0} > 0 \tag{10}$$

To investigate household's integration to the market, a new variable for marketed surplus is defined as

$$M = Q - C_F \tag{11}$$

where M is the total marketed surplus. Differentiating:

$$\frac{\partial M}{\partial \lambda_2} = \underbrace{\frac{\partial Q}{\partial \lambda_2}}_{<0} - \underbrace{\frac{\partial C_F}{\partial \lambda_2}\Big|_{\Delta U=0}}_{>0} - \underbrace{\left(Q - C_F\right)}_{>0} \cdot \underbrace{\frac{\partial C_F}{\partial Y^*}}_{>0} < 0 \tag{12}$$

The marketed surplus of a net seller decreases when the trading costs increase. Not only does the household sell less to the market but it also buys less as shown below. Another measure for market integration is consumption bought from the market (*B*). For a net seller *B* is defined as

$$B = C_{NF} \tag{13}$$

and the impact of increased trading costs is given as

$$\frac{\partial B}{\partial \lambda_1} = \frac{\partial C_{NF}}{\partial \lambda_1} < 0 \tag{14}$$

For the *net buyer*, on the other hand, the income can be re-written in terms of consumption:

$$Y^* = \left(P_F + \lambda_1 - \lambda_2\right) \cdot C_F^* + \left(P_{NF} + \lambda_1\right) C_{NF}^* + W \cdot l^* \tag{15}$$

Increased trading costs imply that the bought food is now more expensive, which encourages the household to buy less from the market and to produce more at home. To see this more formally, differentiating C_F with respect to the trading cost λ_1 yields

$$\frac{\partial C_F}{\partial \lambda_1} = \frac{\partial C_F}{\partial \lambda_1}\Big|_{\Delta Y^*=0} - \underbrace{\frac{\partial C_F}{\partial Y^*} \cdot \frac{\partial Y^*}{\partial \lambda_1}}_{\text{Income effect } < 0} = \underbrace{\frac{\partial C_F}{\partial \lambda_1}\Big|_{\Delta U=0}}_{<0} - \underbrace{\left(C_F - Q\right)}_{>0} \cdot \underbrace{\frac{\partial C_F}{\partial Y^*}}_{>0} < 0 \tag{16}$$

For a net buyer of food, marketed surplus is negative. A more relevant measure of the net buyer's integration to the market is the net consumption bought from the market (*B*) here defined as:

$$B = C_{NF} + \left(C_F - Q\right) \tag{17}$$

where *B* is the amount of goods bought from the market as defined above. Again, differentiating with respect to the trading costs we get:

$$\frac{\partial B}{\partial \lambda_1} = \frac{\partial C_{NF}}{\partial \lambda_1} + \frac{\partial C_F}{\partial \lambda_1} - \frac{\partial Q}{\partial \lambda_1} = \underbrace{\frac{\partial C_{NF}}{\partial \lambda_1}}_{<0} + \underbrace{\frac{\partial C_F}{\partial \lambda_1}\Big|_{\Delta U=0}}_{<0} + \underbrace{\left(Q - C_F\right) \cdot \frac{\partial C_F}{\partial Y^*}}_{<0 \quad >0} - \underbrace{\frac{\partial Q}{\partial \lambda_1}}_{>0} < 0 \tag{18}$$

The net buyer household will withdraw from the market by buying less of both food and non-food products and increasing its domestic production and thus withdrawing its labour from wage labour.

In sum, the optimal response to high trading costs if a household is a net buyer is greater home production and decreased consumption of marketed goods, and conversely high trading costs for the net seller imply reduced selling of the good and more consumption at home. Both sellers and buyers are pushed closer to autarky as transaction costs increase. This highlights the importance of marketing margins on market integration.

3.3 Empirical Application

The theoretical framework discussed in the literature review implies that the households would be better off if they were able to exploit trading opportunities and specialise on what they have comparative advantage of doing. If a household is efficient in farming, it is likely to produce a surplus which it can then sell to increase its consumption of non-food items and leisure. Easiness of trading allows also specialisation within the agricultural sector and thus even the net sellers may end up buying a variety of food items they consume from the market in exchange for their own produce as opposed to producing everything at home. On the other hand, other households may choose to specialise in selling labour and acquiring all food and non-food items from the market. This specialisation would lead to higher welfare levels for both net seller and net buyer households than what can be achieved in autarky. However, the model presented above highlights that transaction costs may hinder the process of specialisation and decrease the benefits of trading and thus lower the welfare of the households. The empirical section of the study aims at measuring the welfare impact gained from trade and specialisation, and identifying the sources of transaction costs and disconnectedness.

As defined above, households' income can be measured in terms of its total consumption. The value of consumption depends on the quantity of goods consumed (including leisure) and their prices. In the model transaction costs affected the value of consumption directly as part of the price in which the consumption was measured, as well as indirectly by affecting the optimal quantity of each good consumed. However, using household specific prices, i.e. consumer prices net of transaction costs, to measure the value of consumption will lead to different measurement for two households consuming identical quantities of all goods if one of the households is a net seller and the other a net buyer. In the empirical analysis the interest is on real consumption across all households and the market prices are used to aggregate the value of consumption of different goods in the consumption basket. In this case the total consumption is a direct function of real quantities consumed and their

market prices, and an indirect function of transaction costs though chosen level of commercialisation, i.e. the choice of optimal resource allocation into agricultural production, wage employment, and allocation of income into different marketed and home produced goods. The household consumption is defined as

$$Y_h = Y_h\left(P^M{}_i \cdot C_{ih}, \gamma_h\right) \qquad \text{h = 1,…,H;} \quad \text{i = 1,…, I} \qquad (19)$$

where Y_h is the total income of household *h*, $P^M{}_i$ is the price of the good *i* at the market, C_{ih} is the quantity of good *i* consumed by household *h*, and γ_h is a measure of commercialisation in household *h*. As discussed above, commercialisation is a summary term of a livelihood strategy choice including labour allocation into home production, buying and selling. As an indicator of households connectedness to the market, consumption bought from the market (*B*) was analysed for both sellers and buyers in the model above. This measure is used in the literature as an indicator of commercialisation. Von Braun and Kennedy (1994, 11-12) suggest using the ratio of bought goods and services over the total income instead of using absolute values in order to make the measure of commercialisation independent of household's initial income. This ratio is also chosen as a proxy for commercialisation in the empirical application.

$$\gamma_h \equiv \frac{B_h{}^M}{Y_h} \qquad (20)$$

where B^M is the value of goods bought from the market measured at market prices. As the theory suggests and the model confirmed, commercialisation depends on transaction costs and income, i.e.

$$\gamma_h = \gamma_h\left(\lambda_h, Y_h\right) \qquad (21)$$

where λ_h is the transaction cost that household *h* faces.

The empirical model aims to find the determinants of household income and determine the importance of market integration in households' welfare. Once the impact of commercialisation on consumption has been measured, exogenous determinants of commercialisation are investigated in

the empirical estimation in order to identify plausible policy instruments for welfare improvement. More detailed description of the estimation methodology and functional forms is presented in section 5.

4 Data

The key dataset used for estimating the empirical model is the Tanzanian Household Budget Survey (HBS) produced by the Tanzanian National Bureau of Statistics (NBS) (NBS 2002a). The Tanzanian HBS 2000/01 collected information from 22,178 households representing the total population of over 35 million. The HBS collected information on a range of individual and household characteristics including household members' education, economic activities and health; household expenditure, consumption and income; ownership of consumer goods and assets; housing structure and materials; and household access to services and facilities. The information was collected using a household questionnaire and daily recordings of household consumption, expenditure and income over a calendar month. (NBS 2002b)

Some statistical characteristics are worth highlighting in order to get a better understanding of the underlying structure of the data. Despite the fact that almost 80% of the population in Tanzania live in rural areas, urban areas are disproportionally represented in the unweighted data forming 65% of the household sample. Given this bias towards easily accessible mostly urban households, also for example the distance to the nearest, but not necessarily most important, markets measure is low for the interviewed households averaging at three kilometres for a one way journey in rural areas and under one kilometre for urban areas. Otherwise, the descriptive characteristics are mostly in accordance to prior hypothesis: the income is lower in the rural areas than in urban areas whereas family size is larger in rural areas compared to the cities. Urban dwellers are also more likely to acquire higher education lever as their rural counterparts. On the whole, the urbanisation has not largely developed in Tanzania and even the urban households report to large extent being farmers: 60% of the rural household heads and 37% of the urban household heads have farming as their main activity in the unweighted sample. However, despite cultivating their own farms urban dwellers are more integrated to the markets and report on average 25 percentage points higher levels for commercialisation than rural households, who still rely on own production for a large part of their consumption. This is not to say that the rural households would not be commercialised as they acquire on average 62% of their consumption from the market. Again, these figures are describing

the unweighted sample which even in rural areas is biased towards more central areas. Wage employment is the dominant source of cash in urban areas whereas agriculture holds the dominant position in the rural areas as the main source of cash. Finally, the sample size is worth highlighting as the survey covered over 20000 households. Such a large dataset is likely to smooth any possible remaining sampling and data errors, and increases the credibility of the obtained results.

5 Methodology

5.1 Consumption

OLS

Estimating total consumption expenditure as presented in equation 19 is done by ordinary least squares (OLS). The OLS equation can be written as,

$$\ln Y_h = X_h\beta + \varepsilon_h \qquad (22)$$

where $\ln Y_h$ is the log of total consumption expenditure in household *h;* $\mathbf{X}_h$ is a vector of household characteristics such as degree of commercialisation, age and education of the household head, household size, main assets (land, cows, sheep) and economic activities; β is the coefficient vector defined as

$$\hat{\beta} = (X'X)^{-1} X'Y \qquad (23)$$

and ε_h is a random error term. The data used for this study was collected in to stages: first sampling primary sampling units (most often villages) and then selecting households from each sampling unit. This implies that the standard errors of the estimation need to be adjusted to take into account that observations within each cluster may not be independent, but that households can be divided into M groups $G_1, G_2, \ldots, G_M$ that are independent. The robust Huber/White/Sandwich estimator[1] is used to obtain robust standard errors in this study as follows:

[1] This estimator was first developed by Huber (1967) and White (1980) who developed the theorem independently of each other.

$$V\left(\hat{\beta}\right) = \left(X'X\right)^{-1} X'\Omega X\left(X'X\right)^{-1} \tag{24}$$

where the Ω is estimated as

$$\hat{\Omega} = \begin{bmatrix} \sigma_{G1}^2 & & & & \\ & \sigma_{G2}^2 & & & \\ & & \ldots & & \\ & & & \sigma_{GM-1}^2 & \\ & & & & \sigma_{GM}^2 \end{bmatrix} \tag{25}$$

Here the values on the diagonal are the variances within each cluster and the values outside the diagonal (covariances between the villages) are all zeros.

Besides clustered sampling, also sample weights were used in the survey to make the results nationally representative. However, including the sampling weights into the actual regression calculation is a matter of debate (see e.g. Deaton 1997, 67-73) and at present no consensus on the matter has been reached. I have chosen not to include the weights into the regressions as is a common practice in many applied econometric studies (see e.g. Heltberg & Tarp 2002).

Endogeneity and Instrumental Variables

As discussed above, there is a case to argue that commercialisation is in fact an endogenous explanatory variable in the consumption equation as wealthier people are more likely to buy from the markets merely because they can afford to do so. If this is true, then estimates using the OLS procedure may be biased. A way to alleviate the problem would be using instrumental variable (IV) estimation. As using the IV technique takes place at the cost of efficiency, the need for IV estimation is tested by using the so called Durbin-Wu-Hausman (DWH) test suggested by Davidson and MacKinnon (1993). The test was first proposed by Durbin (1954) and separately by Wu (1973) and Hausman (1978). Here the null hypothesis is that OLS is unbiased, i.e. any endogeneity among the regressors would not have a deleterious effect on the OLS estimates. The null hypothesis does *not* test whether or not the regressor and the explanatory variable are jointly determined but merely whether it causes the regression estimates to be biased. A rejection of the null indicates that endogenous regressors' effects on the estimates are meaningful and instrumental variables

techniques are required. Under the null, the test variable is distributed Chi-squared with m degrees of freedom, where m is the number of regressors specified as endogenous in the model (Baum et al. 2002). However, Nakamura and Nakamura (1998) point out that the power of the DWH test is positively related to the correlation between the instruments and the endogenous variable, which in cross-section data is often very low. As the instruments become less relevant, the power of the Hausman test decreases and the likelihood of falsly accepting exogeneity increases. Furthermore, they demonstrate that the DWH test is a test of existence of endogeneity, not the severity of such error implying that the DWH test may be significant and yet the OLS bias relatively small or in the case of type two error DWH test can be insignificant and the OLS bias relatively large. Park and Davis (2000) note that if the instruments are weak, the IV estimator will be biased in the same direction as the OLS estimator and the loss of efficiency relative to OLS can be substantial. Following these words of warning a practical advice of Nakamura and Nakamura is accepted and both IV and OLS results are reported in this study.

In order to test for the endogeneity and indeed to correct the estimates if needed, one needs to find instruments that are correlated with the endogenous variable, excluded from the main equation, and orthogonal to the disturbance process. The instruments to be used to test the endogeneity of commercialisation in the consumption equation are variables correlated with commercialisation but not directly with consumption. The distance to the nearest market is selected as the first instrument, as it is assumed to be a significant determinant of households' decision to trade. However, it is unlikely that distance to the market has a significant direct effect on income over and above the indirect effect through commercialisation bearing in mind that the consumption (including own production) is measured at market prices. Distance will then not affect the income received from selling goods, or decrease the value of own production. It will, however, impact the buying and selling behaviour and this is where the indirect effect takes place. Distance can also be argued to be exogenous to the households considering the country specific circumstances in Tanzania, where tenure is traditionally long-lived. The land has usually been originally allocated by the chiefs and local leaders to each family and clan, who can then pass on the right to cultivate the land to their children (Mtetewaunga 1986). Most of the land is inherited from the parents and households are unlikely to move from the neighbourhood where they were born. In the unlikely event that a wage employment opportunity raises, individual family members may even move to the other side of the country, but most often the inherited land offers the main source of income and food security.

Despite the fact that formally speaking all land belongs to the state, and that legislation has been changed gradually to implement the heavily debated land reform programme aimed at transferring the ownership of land to the private sector, buying and selling land is still rather unusual in Tanzania. As concluded by Wily (2000) "The facts are these: despite a century of purposeful penetration by non-customary tenure ideology and legal provision, unregistered, customary tenure not only persists but is still by far the majority form of tenure in the region. None of the strategies adopted to ignore or diminish it have been successful." This evidence from previous studies supports the view that distance to the market is exogenous determinant of market integration.

Another instrument correlated only with commercialisation is related to market information and ties in to the same discussion of low social mobility. Even though informal sources of information are still dominant especially in rural Tanzania, the provision of market information is generally poor and news do not travel fast between the villages. People close to the market are likely to get relevant information directly from the market, but complementary sources of market information are the media (radio and TV) and telephone. These productive assets are used to gain information on the situation in other markets, likely supply and demand at the nearby market, and to gain other information on which to base the trading decisions on. The ownership of a radio, TV or a telephone has been used as a proxy for the household's access to market information previously in the literature (see e.g. Heltberg and Tarp 2002) and the same assumption is used in this study. Even though the total value of the household assets is likely to have direct impact on household's income, once the overall value of assets is controlled for, the type of assets that the household owns can be used as an indicator of the kind of productive resources that the household has available to them. Both of the above presented instruments are included as dummy variables, and they have also been interacted with the regional dummy variables to capture the differences of the effect in different areas of the country (e.g. the significance of distance in mountainous versus plain area).

The task of selecting the instruments is a particularly important one as the reliability of the results relies on the validity of the instruments. Besides the theoretical justification discussed above, the instruments need to work for the dataset given. The Wu-Hausman test is vulnerable to the endogeneity of the instruments as it cannot differentiate between an exogenous explanatory variable and an endogenous instrument, on one hand, and an endogenous explanatory variable and exogenous instrument, on the other. Statistical testing of the instruments used is therefore and

prerequisite for the use of the Wu-Hausman test. A commonly used test for instrument validity is the so called Sargan test of overidentifying restrictions proposed by Sargan (1958). The joint null hypothesis of the test is that the excluded instruments are valid, i.e. that they are uncorrelated with the error term, and that they are correctly excluded from the estimated equation. Under the null, the test statistics is distributed as chi-squared in the number of overidentification restrictions, i.e. the exogenous instruments. However, as the data was collected using primary sampling units, the clustering needs to be taking into account also when calculating the test statistics. As the Sargan statistics is not consistent in presence of heteroscedasticity, Hansen's J statistic (Hansen 1982) is reported instead as it is robust to heteroscedasticity of unknown form. These test statistics are closely linked as, indeed, under conditional homoscedasticity, Hansen's J becomes Sargan's statistics (Hayashi 2000, pp. 227-228), and thus Hansen's J is also known as Hansen-Sargan test statistics. The test statistics is specified as follows:

$$J\left(\hat{\delta}\left(\hat{S}^{-1}\right),\hat{S}^{-1}\right)=n\cdot g_n\left(\hat{\delta}\left(\hat{S}^{-1}\right)\right)\hat{S}^{-1}g_n\left(\hat{\delta}\left(\hat{S}^{-1}\right)\right)\rightarrow\chi^2\left(K-L\right) \qquad (26)$$

where n is the sample size, $g_n\left(\hat{\delta}\left(\hat{S}^{-1}\right)\right)$ are the ortoghonality or moment conditions, and $\hat{S}^{-1}$ is the optimal weighting matrix (Hayashi 2000, p. 217). The estimator uses a cluster-robust optimal weighting matrix and the estimator is robust to arbitrary intra-cluster correlation.

The test is based on a normal IV estimation with a continuous dependent variable and a normally distributed error term. However, as mentioned the commercialisation variable is defined as a proportion of the consumption that is bought from the market, i.e. as a proportion bound between 0 and 1. Using OLS to create the predicted values for commercialisation is likely to lead predictions over 1 or under 0. Still, this does not cause problems when testing the instrument validity for the consumption equation as the IV is a single equation method and how the instrumented variable is modelled is irrelevant for the consistency. The test results of the Hansen-Sargan test are reported for all the equations together with the other estimation results.

The possibility of endogeneity has been taken into account also in the estimation procedure and instrumental variable regression is used as an alternative way of estimating the consumption equation. Normally the endogenous variable (here: commercialisation) is first regressed on all exogenous variables in the system, and then the predicted values are used instead of the original

variable in the estimation. However, strictly speaking as commercialisation is not a linear function of its regressors, the estimates should not be used directly to replace commercialisation in the consumption equation as this would lead to invalid OLS standard errors, the consistency of the estimates would rely heavily on correctly specified model for commercialisation, and asymptotic variance of the estimator might be biased (Wooldridge 2002, pp. 623-625). Instead Wooldridge proposes estimating the commercialisation equation first, then taking the predicted values and using these as an exogenous instrument for commercialisation instead of using the original instruments directly. Even though the modified estimation procedure makes little difference to the obtained estimates, now the usual standard errors and test statistics are asymptotically valid. This specification has also an important robustness property as when using the predicted values of commercialisation as instrument, the model which gave the predictions does not have to be correctly specified. In this case the estimator for β is given by

$$\hat{\beta}_{IV} = (W'X)^{-1}W'Y \tag{27}$$

where W is the n x k matrix of instruments and here k =1. The robust Huber-White variance estimator is then

$$V(\beta_{IV}) = (W'X)^{-1}W'\Omega W(W'X)^{-1} \tag{28}$$

5.2 Commercialisation

Explaining the determinants of commercialisation presents interesting methodological issues to be considered as here the dependent variable is a fraction of the household's total consumption acquired through market transactions. Thus, the values of commercialisation are bound between 0 and 1 so that $0 \le \gamma_h \le 1$ where the extreme value 1 is reasonably common among wealthier urban dwellers. The standard least squares model is not equipped to cope with such limitations of the feasible values for the dependent variable. A commonly used replacement for cases when the dependent variable is bound between zero and one is the standard logit function $\log\left[\gamma_h / (1-\gamma_h)\right]$. However, this is not defined if γ_h takes on the extreme values 0 or 1 with positive probability. Consequently in any dataset where an observation γ_h equals 0 or 1 an adjustment must be made before computing the log-odds ratio. Commonly used but methodologically unsatisfactory

adjustment is to replace 0 and 1 with an arbitrary value very close to these extreme values. Especially when a large proportion of the observations take on the extreme values in the dataset used, adjusting the original observations is not desirable.

As 16% of the unweighted observations in the Tanzanian sample reported buying all the goods they consumed from the market placing an arbitrary value instead of value 1 to be able to use the standard logistic function is not appropriate for the model. Instead in this study a Generalised Linear Model (GLM) approach presented by Papke and Wooldridge (1996) is used in the commercialisation equation to take into account the characteristics of the proportional dependent variable where the logit transformation is made for the *expected* value of the dependent variable[2]. For the estimation the functional form of the expected rate of commercialisation of household h, γ_h conditional on household characteristics $\mathbf{Z_h}$[3] is defined as

$$E\left(\gamma_h \middle| \mathbf{Z}_h\right) = G\left(\mathbf{Z}_h\beta\right) \tag{29}$$

where $\mathbf{Z_h}$ is a vector of household specific characteristics including distance to the nearest market and access to market information; household demographics (such as the household size, sex of the household head, education of the household head); information about household's economic activities and assets (such as ownership of land, cattle, total assets); as well as regional and seasonal controls. $G(\cdot)$ is the link function which here is defined as a cumulative distribution function that is assumed to be the logistic function

$$G\left(\mathbf{Z}_h\beta\right) = \frac{\exp\left(Z_h\beta\right)}{1+\exp\left(Z_h\beta\right)} \tag{30}$$

[2] For other applications of the methodology see e.g. Hausman & Leonard (1997).

[3] The possibility of including consumption as an endogenous explanatory variable was also explored by including consumption and instrumented values for consumption in the commercialisation equation. The results were found to be very similar to but less robust than the specification where consumption is not included. As the theory does not suggest that consumption would directly cause commercialisation but the causality is likely to be reversed, and as the aim of the exercise is to identify plausible policy measures to enhance consumption through commercialisation, consumption was excluded from the list of explanatory variables in the commercialisation equation.

In other words, instead of taking the standard logit transformation of the dependent variable, i.e. $E\left(\ln\left(\frac{\gamma_h}{1-\gamma_h}\right)\middle|Z_h\right)=Z_h\beta$, the current approach transforms the expected value of the estimate instead, i.e. $E\left(\gamma_h\middle|Z_h\right)=\frac{\exp\left(Z_h\beta\right)}{1+\exp\left(Z_h\beta\right)}$. The latter is always defined whereas the former is not if the dependent variable takes on values 1 or 0. The parameters in (29) can be estimated using maximum likelihood technique where the likelihood for an observation is specified as the Bernoulli log-likelihood, i.e.

$$L_h=\gamma_h\ln\left(\frac{\exp\left(Z_h\beta\right)}{1+\exp\left(Z_h\beta\right)}\right)+\left(1-\gamma_h\right)\ln\left[1-\frac{\exp\left(Z_h\beta\right)}{1+\exp\left(Z_h\beta\right)}\right] \tag{31}$$

These estimates are consistent as long as the conditional expectation (29) is correctly specified even if the Bernoulli specification is incorrect (Papke and Wooldridge 1996). The maximisation problem is then

$$\max_{\beta}\sum_{h=1}^{H}L_h\left(\mathbf{Z}_h\beta\right) \tag{32}$$

As mentioned, the households have been selected in two stages by first selecting the sampling unit and then households to be interviewed within the sample. The possibility that the sampling errors are correlated within the sampling unit but not across the units is taken into account when calculating the variances by using the robust sandwich estimator. Here the variance is defined as

$$\hat{\nu}=\hat{V}\left(\sum_{k=1}^{M}u_k^{(G)'}u_k^{(G)}\right)\hat{V} \tag{33}$$

where $\hat{V}=\left(\frac{-\partial^2\ln L}{\partial\beta^2}\right)^{-1}$ i.e. the conventional estimator of variance, which here is the inverse of the estimated information matrix. This matrix has been adjusted by the score weights for each cluster

group. $u_k^{(G)}$ is the contribution of the *k*th group to the scores $\left(\frac{\partial \ln L}{\partial \beta}\right)^2$. (Rogers 1993, Williams 2000, Wooldridge 2002) This formula is asymptotically correct, but as the sample is not infinite, the matrix is multiplied by $\frac{M}{M-1}$ for finite sample correction, even though in such a large sample this correction does not have large practical significance. The resulting quasi-maximum likelihood model is estimated using Newton-Raphson optimisation.

6 Results

6.1 Households' total consumption

As described above, the estimation of the total consumption per adult equivalent was done by using OLS and IV with a correction for robust standard errors. The estimation results are presented below in table 1. Whether or not the endogeneity bias was in fact present was tested for all model specifications. The results suggest that in most cases there was no endogeneity bias present. On the other hand, the correction for possible endogeneity had little impact on the estimations as a whole and the coefficients for the instrumented variables turned out to be very similar to the results from pure commercialisation. In fact in some cases the coefficient for the instrumented variable was even higher than for pure commercialisation as the instruments, such as distance to the markets, are likely to capture a wider range of elements determining income generation that are related to isolation from the market and transaction costs. Only for the urban sub-sample the coefficient estimate for the commercialisation falls substantially when instrumented variables are used instead of the actual ones (see appendix). This does not undermine the importance of commercialisation in urban areas but is likely to imply that the instruments used, i.e. distance to the market and access to market information, are not as powerful elements to determine market integration in urban areas as they are in rural areas where distances are longer and access poses a real problem for the households. The instruments generally passed the Hansen-Sargan test for instrumental validity which tests whether or not the instruments are orthogonal to the error process but is silent on the impact of the instruments. On the whole, the estimates for commercialisation remain very robust through all specifications.

Furthermore, other possibly endogenous variables that are important contributors to welfare are sources of cash income and employment category. Even though these characteristics are likely to determine household's welfare, it may also be that their occupational choice is related to income, e.g. only rich households sell agricultural output as they can afford to do so. The model is estimated and the results are presented with and without these possibly endogenous categories. The exclusion of the variables has practically no impact on the other coefficients but it worsens the fit of the model, as was to be expected. Finally, the model is estimated over the whole sample as well as separately for urban and rural areas, since the pooling test of joint significance of all interaction terms indicate that a simple pooled model is not sufficient. The results for the rural and urban samples separately are included as an appendix.

On the whole, the model seems to fit the data well as around 20 percent of the variation in the household total expenditure can be explained in all the models. Also Ramsey's RESET is passed in all full models at the conventional five percent confidence level and thus the null hypothesis of no omitted variables or incorrect functional form is maintained. The reduced form of the equation commonly fails the RESET for known reasons: important explanatory variables such as employment category, have been deliberately omitted to demonstrate the robustness of other results.

TABLE 1: COEFFICIENT ESTIMATES FOR HOUSEHOLD TOTAL CONSUMPTION PER ADULT EQUIVALENT FOR 28 DAYS (POOLED MODEL)

	Reduced OLS	Reduced IV	Full OLS	Full IV
Commercialisation	0.489***	-	0.408***	-
	(13.46)	-	(11.03)	-
Predicted values	-	0.476***	-	0.303*
	-	(3.05)	-	(1.93)
HOUSEHOLD DEMOGRAPHICS				
HH has only 1 member	†	†	†	†
	(.)	(.)	(.)	(.)
HH has 2-6 family members	-0.527***	-0.527***	-0.543***	-0.543***
	(31.07)	(31.00)	(32.54)	(32.50)
HH has 7 or more members	-0.895***	-0.897***	-0.925***	-0.926***
	(44.94)	(44.98)	(47.17)	(47.29)
HH head is male	-0.009	-0.009	-0.009	-0.010
	(0.82)	(0.80)	(0.85)	(0.90)
HH head age	-0.001**	-0.001***	-0.001***	-0.001***
	(2.56)	(2.81)	(2.61)	(2.86)
HH head sick in last 4 weeks	0.030***	0.030***	0.032***	0.031***
	(2.90)	(2.83)	(3.11)	(3.04)
HH head has no education	†	†	†	†
	(.)	(.)	(.)	(.)
HH head has primary education	0.018	0.016	0.020	0.019
	(1.31)	(1.13)	(1.49)	(1.43)
HH head has secondary educ.	0.025	0.024	0.028	0.029
	(1.30)	(1.20)	(1.46)	(1.48)
HH head has higher education	0.056**	0.055**	0.061***	0.061***
	(2.49)	(2.44)	(2.76)	(2.73)
MAIN ACTIVITY IS FARMING				
The main activity of the HH	0.011	0.012	0.014	0.013
head is farming/livestock	(0.89)	(0.92)	(1.11)	(1.04)
PRODUCTIVE ASSETS				
log of assets	0.002	0.003	0.002	0.003
	(0.62)	(0.79)	(0.68)	(0.94)
HH has no land	†	†	†	†
	(.)	(.)	(.)	(.)
HH owns max 2 acres of land	-0.051***	-0.051***	-0.037**	-0.041**
	(3.40)	(2.85)	(2.47)	(2.46)
HH owns 2-10 acres of land	-0.015	-0.015	0.003	-0.002
	(0.85)	(0.70)	(0.17)	(0.09)
HH owns more than 10 acres	0.056**	0.052*	0.088***	0.083***
	(2.09)	(1.84)	(3.30)	(3.08)
HH has 1-10 cows	0.081***	0.080***	0.078***	0.077***
	(4.99)	(4.93)	(5.00)	(4.93)
HH has more than 10 cows	0.056*	0.059*	0.067**	0.070**
	(1.74)	(1.81)	(2.11)	(2.20)
HH has 1-10 sheep	0.040***	0.038***	0.029**	0.026*
	(2.87)	(2.59)	(2.10)	(1.77)
HH has more than 10 sheep	0.073***	0.075***	0.061***	0.060***
	(3.27)	(3.28)	(2.73)	(2.67)
HH has access to electricity	0.295***	0.295***	0.278***	0.281***
	(17.67)	(15.81)	(17.19)	(16.34)
INCOME SOURCES				
HH sells agricultural output	-	-	0.120***	0.116***
			(8.75)	(7.88)
HH sells non-farm output	-	-	0.119***	0.121***

			(10.55)	(9.75)
MAIN SOURCE OF CASH				
HH's main source of cash is agriculture	† (.)	† (.)	† (.)	† (.)
HH's main source of cash is wage or business income	- -	- -	0.165*** (11.04)	0.176*** (8.74)
HH's main source of cash is remittances or other undefined	- -	- -	-0.023 (1.29)	-0.018 (0.99)
SEASONAL VARIATION				
January	† (.)	† (.)	† (.)	† (.)
February	0.027 (1.37)	0.026 (1.36)	0.029 (1.51)	0.027 (1.42)
March	0.010 (0.49)	0.011 (0.52)	0.012 (0.56)	0.011 (0.54)
April	0.001 (0.05)	-0.002 (0.08)	-0.000 (0.00)	-0.003 (0.16)
May	-0.006 (0.28)	-0.006 (0.27)	-0.007 (0.34)	-0.008 (0.37)
June	-0.034 (1.58)	-0.031 (1.45)	-0.037* (1.74)	-0.035* (1.65)
July	0.036* (1.74)	0.035* (1.68)	0.029 (1.42)	0.028 (1.37)
August	0.021 (1.02)	0.021 (1.00)	0.012 (0.61)	0.011 (0.54)
September	0.034* (1.66)	0.035* (1.71)	0.029 (1.43)	0.030 (1.50)
October	-0.017 (0.87)	-0.018 (0.93)	-0.025 (1.30)	-0.026 (1.31)
November	0.027 (1.39)	0.029 (1.45)	0.022 (1.16)	0.024 (1.21)
December	0.117*** (5.61)	0.117*** (5.57)	0.122*** (5.88)	0.123*** (5.91)
REGIONAL VARIATION				
HH lives in urban area	0.021 (0.97)	0.025 (0.67)	0.014 (0.63)	0.031 (0.92)
Dodoma	† (.)	† (.)	† (.)	† (.)
Arusha	-0.160*** (3.72)	-0.162*** (3.61)	-0.125*** (2.89)	-0.120*** (2.64)
Kilimanjaro	-0.094** (2.30)	-0.090** (1.97)	-0.047 (1.16)	-0.032 (0.68)
Tanga	-0.014 (0.34)	-0.011 (0.25)	0.029 (0.71)	0.040 (0.92)
Morogoro	0.089** (1.99)	0.091** (1.99)	0.129*** (2.96)	0.135*** (3.04)
Pwani	-0.035 (0.78)	-0.032 (0.65)	-0.012 (0.27)	0.002 (0.04)
Dar es Salaam	0.037 (0.80)	0.039 (0.83)	0.076* (1.68)	0.085* (1.81)
Lindi	0.031 (0.54)	0.033 (0.57)	0.060 (1.04)	0.067 (1.15)
Mtwara	-0.024 (0.40)	-0.020 (0.34)	-0.004 (0.07)	0.004 (0.06)
Ruvuma	-0.026 (0.53)	-0.019 (0.39)	-0.010 (0.22)	-0.003 (0.07)
Iringa	0.168***	0.164***	0.185***	0.182***

	(3.19)	(3.12)	(3.68)	(3.64)
Mbeya	0.121***	0.121***	0.169***	0.172***
	(2.89)	(2.88)	(4.01)	(4.04)
Singida	-0.113***	-0.102**	-0.100**	-0.089**
	(2.63)	(2.43)	(2.34)	(2.13)
Tabora	0.169***	0.173***	0.204***	0.216***
	(3.46)	(3.41)	(4.38)	(4.42)
Rukwa	-0.097**	-0.095**	-0.063	-0.058
	(2.08)	(2.04)	(1.38)	(1.27)
Kigoma	0.102**	0.105**	0.114**	0.122**
	(2.10)	(2.11)	(2.38)	(2.48)
Shinyanga	0.224***	0.228***	0.247***	0.250***
	(5.18)	(5.25)	(5.78)	(5.82)
Kagera	0.232***	0.234***	0.247***	0.250***
	(5.48)	(5.55)	(5.78)	(5.86)
Mwanza	-0.069	-0.075	-0.017	-0.009
	(1.43)	(1.41)	(0.35)	(0.18)
Mara	-0.284***	-0.280***	-0.228***	-0.207***
	(4.53)	(4.27)	(3.73)	(3.19)
Constant	9.435***	9.440***	9.328***	9.383***
	(148.60)	(87.77)	(148.63)	(91.24)
Observations	21723	21530	21723	21530
R-squared	0.24	0.20	0.26	0.16
RESET test	5.30***	3.37**	1.35	0.14
Prob > F	0.001	0.018	0.257	0.935
Hansen-Sargan	-	89.203	-	88.163
Prob > Chi-sq	-	0.161	-	0.181
Endogeneity test	0.00	-	0.44	-
Prob > F	0.947	-	0.509	-
Pooling test	2.05***	95.09***	1.85***	87.65***
Prob > F	0.000	0.000	0.000	0.000

Robust t statistics in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%
† denotes base category

As the theory suggests and these results confirm commercialisation has a significant positive effect on consumption. A one percentage point increase in commercialisation leads to, on average and *ceteris paribus*, 0.3-0.5 percentage increase in the consumption in the pooled sample, 0.5-0.7 percentage increase in rural areas, and 0.1-0.4 percentage increase in urban areas, respectively. This is a sizable effect which implies that commercialisation has a strong potential as a tool in the fight against poverty especially in the rural areas where the households face largest barriers for trade. The current average commercialisation rate in Tanzania is 63 percent when the sample is adjusted for the total population and even though none of the households lives in total subsistence, there is considerable scope for improvement.

On average, households with only one member consume more per capita as do larger families. This result was to be expected as households usually share durable goods, which are counted only once for the whole household. However, the larger the family size is, the lower the per capita expenditure, which implies that not only are the common goods shared with larger number of people, but the consumption of individual goods is also lower per person. This empirical result is commonly found in literature (see e.g. Lipton & Ravallion 1995, Deaton and Paxton 1998) even though it is in contrast with the economic theory of economies of households due to sharing of household public goods (Lanjouw & Ravallion 1995). The majority of empirical findings from different countries still show that larger households consume less per capita, which is the case also in Tanzania. The sex of the household head is not statistically significant determinant of household consumption in the sample. Even though there are far fewer households headed by women, controlling for all other explanatory factors the sex of the household head alone does not have an effect on household income. On the other hand the age of the household head is a well determined factor of consumption implying that the older the head of the household, the smaller the average consumption of a household member. The age effect is still very small, on average 0.1 percent per year, and thus the lower consumption due to the head of the household mostly impacts households headed by over 60 year-olds (13% of the sample) who might no longer be as efficient in acquiring income for the household as the younger household heads may be, and the older household heads are also more likely to have larger extended households to look after. The oldest member of the household is still often selected as the household head even though he/she might not be the most educated or best equipped for making decisions of using the household resources. Somewhat counter intuitively, if the household head has been ill or injured during the period when the

consumption has been measured, the total expenditure seems to be larger than otherwise. Even though well determined, the sickness only increases the consumption by three percent in the pooled sample, one percent in the rural sample and four percent in the urban sample. This might be caused by the household's attempt to nurture the household head back into health by borrowing or depleting stocks in order to get him/her back into more productive activities as soon as possible. Rural households might not have as much leverage to increase their consumption as the urban households have in times of crises and thus their consumption is close to unaffected. Education, on the other hand, has the expected positive effect on total consumption. The more educated the household head is, the higher the average consumption of the household members, even though the impact of basic education is rather small, controlling for other factors. If the household head has higher education after the secondary school, the total consumption of a household member is six percent greater compared to a household head with no education in the pooled sample. In the rural areas where educated people are on short supply returns to education are even greater: A household head with secondary education in rural areas consume, on average and *ceteris paribus*, 10-11 percent more than a household whose head has no education.

Economic activities of the household are further possible determinants of consumption. Whether or not the household head is a farmer or not does not have a statistically significant impact on consumption most likely as farming is so common across both rich and poor groups in rural and urban Tanzania. However, in the full models where other information on economic activities are included a household that engages in selling agricultural output consumes, on average and *ceteris paribus*, 11-13 percent more than a household that does not engage in trade in all samples. The same applies to households that sell non-farm output where the welfare gain is 13, 9, and 15 percent if a household sells non-agricultural products in pooled, rural and urban samples, respectively. Thus whether or not the household engages in selling matters and the gains are sizable irrespective of the goods sold. On the employment side, if household's main source of cash is wage, salary or other business income the household consumption is 18-20 percent higher in all models on average and *ceteris paribus* compared to the households where the main source of cash is from agriculture. Thus employment opportunities, often associated with further commercialisation, are important determinants of income. On the other hand, households depending on remittances, transfers and other casual cash earning opportunities are worse off than households who have more reliable cash income from agriculture, even though this impact is not statistically significant.

The type of assets owned seems to matter more than the total value of the assets in terms of household's total consumption. Households owning only small plots of land are worse off than households without any land, which are often urban dwellers. Still, owning larger plots of land has a significant positive impact on consumption in the rural areas as was to be expected. Also owning cattle and sheep are positively related to consumption, cattle more so in the rural areas and sheep in the urban areas. Especially cattle are often used not only as help in cultivation but also as a means of investment and an indicator of wealth in Tanzania.

Consumption does not vary significantly over different calendar months apart from a peak in December. This is explained by the holiday season which boosts the average consumption by 12-13 percent on average in the whole country, 6-7 percent in rural areas and 13-14 percent in urban areas, compared to January, which was used as a base. Regional variation, on the other hand, is much larger across the regions. Also living in an urban area as such, after all other explanatory factors have been controlled for, has a positive but not statistically significant impact on consumption. This implies that even though consumption levels are lower in rural areas, these differences are explained by other factors, such as employment opportunities, integration to the markets and access to education, rather than living in rural areas as such. The regional comparison category was Dodoma in central Tanzania, which represents a considerably sizable region with average mean income levels including both rural and urban areas. As was found in the final report of the Household Budget Survey 2000/01 (NBS 2002b) consumption expenditure per capita varies considerably across the regions and these differences are not washed away by the other explanatory variables in the model. Living in the capital area has an expected positive and significant impact on consumption of urban population boosting it by 10-15 percent on average and *ceteris paribus*.

Disaggregating the results even further reveals, that the difference in the consumption levels in the rural and urban areas is largely due to the endowment effect, i.e. rural households make as effective use of their resources as urban households but their worse-off position is due to lack of endowments. The most sizable endowment effect comes from wage earning possibility and access to electricity, both of which are much more common in urban areas and have sizable effect on total consumption. Statistically the most significant difference in the treatment is for land, cattle and sheep as was to be expected, but also commercialisation and selling non-farm output had a significant interaction effect with the urban dummy which implies that different types of

commercialisation efforts have different impact in rural and urban areas, as seen in the estimation results above.

6.3 Commercialisation

As described in section 5.2 commercialisation equation was modelled using generalised linear model (GLM) technique, which was estimated using the Newton-Raphson maximum likelihood estimation. As above, the results are presented with and without the employment categories that might be affected by commercialisation. The results were calculated firstly for the whole sample and then separately for the rural and the urban sub-samples as splitting the sample was suggested by the likelihood ratio test. The resulting models fit the data well and are able to explain around 40 percent of the variation in the observed commercialisation over the whole sample, and around 20-30 percent in the sub-samples. Furthermore, even though the RESET for omitted variables or incorrect functional form is failed for the pooled sample, the test is passed in the urban and rural sub-samples at the conventional five percent level, as they succeed better in explaining commercialisation than the pooled sample alone. Still, the RESET is based on OLS estimation, and thus the results of the test should be interpreted with caution. The estimation results are presented in the table 2. As log odds ratios are not intuitively appealing, the results are transformed into impact effects of each dummy variable. The table presents the predicted commercialisation for an imaginary household with mean characteristics on the continuous variable ‘assets’ and for whom all dummy variables take the value of zero. This base is constructed to be highly disadvantaged rural household with no education, land, cattle, sheep, or radio; who earn most of their cash income from agriculture; lives in the base category region Dodoma and is interviewed in the base month January. Each dummy variable is then introduced one at the time and the new predicted rate of commercialisation, as well as the deviation from the base, is calculated to show explicitly the impact of each dummy variable. The deviations from the base occur singly.

Table 2: Commercialisation rates[a] by Household characteristics

	POOLED Reduced	POOLED Full	RURAL Reduced	RURAL Full	URBAN Reduced	URBAN Full
Base[b]	0.72	0.66	0.63	0.56	0.88	0.84
Deviation[c]	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ACCESS TO MARKET INFORMATION						
Radio, TV or telephone	0.73 (0.01)***	0.68 (0.01)**	0.64 (0.01)*	0.57 (0.01)	0.89 (0.01)***	0.85 (0.01)**
DISTANCE TO THE NEAREST MARKET						
within 1 km	† (.)	† (.)	† (.)	† (.)	† (.)	† (.)
1 km	0.72 (0.00)	0.66 (0.00)	0.62 (-0.01)	0.55 (0.00)	0.88 (0.00)	0.85 (0.01)
2 km	0.68 (-0.04)***	0.63 (-0.03)***	0.59 (-0.04)***	0.53 (-0.03)***	0.86 (-0.02)**	0.82 (-0.02)**
3 km	0.67 (-0.05)***	0.61 (-0.05)***	0.57 (-0.06)***	0.50 (-0.06)***	0.85 (-0.02)***	0.82 (-0.02)**
4 km	0.68 (-0.04)***	0.64 (-0.03)**	0.58 (-0.05)***	0.52 (-0.04)**	0.87 (-0.01)	0.83 (-0.01)
5 km	0.68 (-0.04)**	0.64 (-0.03)	0.60 (-0.03)	0.53 (-0.02)	0.81 (-0.07)***	0.78 (-0.06)**
6 km	0.63 (-0.09)***	0.58 (-0.08)***	0.54 (-0.09)***	0.48 (-0.08)***	0.82 (-0.06)**	0.78 (-0.06)*
7 km	0.71 (-0.01)	0.66 (-0.01)	0.62 (-0.01)	0.55 (0.00)	0.91 (0.03)	0.89 (0.05)
8 km	0.66 (-0.06)**	0.61 (-0.05)	0.57 (-0.06)*	0.52 (-0.04)	0.73 (-0.15)**	0.64 (-0.19)**
9 km	0.64 (-0.08)**	0.59 (-0.07)*	0.55 (-0.08)**	0.50 (-0.06)	0.86 (-0.02)	0.80 (-0.04)
10 km or more	0.64 (-0.08)***	0.60 (-0.07)***	0.56 (-0.07)***	0.50 (-0.06)***	0.80 (-0.08)***	0.78 (-0.06)**
HOUSEHOLD DEMOGRAPHICS						
only 1 member	† (.)	† (.)	† (.)	† (.)	† (.)	† (.)
2-6 family members	0.73 (0.01)	0.67 (0.01)	0.60 (-0.03)**	0.53 (-0.03)*	0.90 (0.02)***	0.86 (0.02)***
7 or more members	0.72 (0.00)	0.66 (0.00)	0.59 (-0.04)***	0.52 (-0.04)***	0.90 (0.02)***	0.86 (0.02)**
HH head is male	0.71 (-0.01)*	0.66 (-0.01)	0.62 (-0.01)	0.55 (-0.01)	0.87 (-0.01)	0.83 (-0.01)
HH head has no education	† (.)	† (.)	† (.)	† (.)	† (.)	† (.)
HH head has primary educ.	0.74 (0.02)***	0.68 (0.02)***	0.64 (0.01)	0.57 (0.01)*	0.89 (0.01)***	0.85 (0.01)***
HH head has secondary educ.	0.75 (0.03)***	0.69 (0.03)***	0.64 (0.01)	0.57 (0.01)	0.90 (0.02)***	0.87 (0.03)***
HH head has higher educ.	0.72 (0.00)	0.67 (0.00)	0.63 (0.00)	0.56 (0.01)	0.88 (0.00)	0.84 (0.00)
FARMING/LIVESTOCK AS THE MAIN ACTIVITY						
Main activity is farming	0.70 (-0.02)***	0.65 (-0.02)***	0.62 (-0.01)	0.55 (-0.01)	0.86 (-0.02)***	0.82 (-0.02)***
PRODUCTIVE ASSETS						
HH has no land	† (.)	† (.)	† (.)	† (.)	† (.)	† (.)
HH owns max	0.60	0.56	0.54	0.48	0.80	0.77

2 acres of land	(-0.12)***	(-0.10)***	(-0.09)***	(-0.08)***	(-0.08)***	(-0.07)***
HH owns 2-10	0.57	0.55	0.52	0.48	0.77	0.75
acres of land	(-0.15)***	(-0.11)***	(-0.11)***	(0.08)***	(-0.11)***	(-0.09)***
HH owns > 10	0.60	0.59	0.55	0.50	0.80	0.79
acres of land	(-0.12)***	(-0.08)***	(-0.08)***	(-0.05)***	(-0.08)***	(-0.05)***
1-10 cows	0.71	0.66	0.63	0.56	0.86	0.83
	(-0.01)*	(0.00)	(0.00)	(0.01)	(-0.02)***	(-0.01)**
> 10 cows	0.69	0.65	0.60	0.55	0.87	0.84
	(-0.03)**	(-0.01)	(-0.03)*	(-0.01)	(-0.01)	(0.00)
1-10 sheep	0.69	0.63	0.62	0.55	0.84	0.80
	(-0.03)***	(-0.03)***	(-0.01)	(-0.01)	(-0.04)***	(-0.04)***
> 10 sheep	0.70	0.64	0.64	0.56	0.85	0.81
	(-0.02)*	(-0.02)**	(0.01)	(0.00)	(-0.03)***	(-0.03)***
HH has access	0.81	0.74	0.76	0.67	0.92	0.88
to electricity	(0.09)***	(0.07)***	(0.13)***	(0.11)***	(0.04)***	(0.04)***
INCOME SOURCES						
HH sells	-	0.62	-	0.54	-	0.78
agricultural output		(-0.05)***	-	(-0.02)**	-	(-0.06)***
HH sells non-	-	0.71	-	0.63	-	0.86
farm output	-	(0.05)***	-	(0.07)***	-	(0.02)***
MAIN SOURCES OF CASH						
agriculture	†	†	†	†	†	†
	(.)	(.)	(.)	(.)	(.)	(.)
wage or	-	0.78	-	0.69	-	0.90
business income-		(0.11)***	-	(0.13)***	-	(0.06)***
remittances or	-	0.66	-	0.55	-	0.83
other undefined	-	(0.00)	-	(-0.01)	-	(-0.01)
SEASONAL VARIATION						
January	†	†	†	†	†	†
	(.)	(.)	(.)	(.)	(.)	(.)
February	0.70	0.64	0.63	0.54	0.87	0.82
	(-0.02)**	(-0.02)**	(0.00)	(-0.02)	(-0.01)**	(-0.02)**
March	0.70	0.63	0.61	0.53	0.86	0.82
	(-0.02)***	(-0.03)***	(-0.02)	(-0.03)	(-0.01)**	(-0.02)**
April	0.70	0.65	0.62	0.55	0.87	0.83
	(-0.02)	(-0.02)	(-0.01)	(-0.01)	(-0.01)	(-0.01)
May	0.70	0.64	0.68	0.60	0.85	0.81
	(-0.02)	(-0.02)*	(0.05)**	(0.05)*	(-0.03)***	(-0.03)***
June	0.71	0.65	0.68	0.60	0.86	0.82
	(-0.01)	(-0.01)	(0.05)**	(0.04)**	(-0.02)**	(-0.02)**
July	0.72	0.66	0.70	0.62	0.85	0.81
	(0.00)	(0.00)	(0.07)***	(0.06)***	(-0.02)***	(-0.02)***
August	0.71	0.65	0.69	0.61	0.85	0.81
	(-0.01)	(-0.01)	(0.06)***	(0.05)**	(-0.03)***	(-0.03)***
September	0.73	0.67	0.71	0.63	0.87	0.82
	(0.01)	(0.01)	(0.08)***	(0.07)***	(-0.01)*	(-0.01)*
October	0.74	0.68	0.71	0.63	0.88	0.84
	(0.02)*	(0.02)	(0.08)***	(0.07)***	(0.00)	(0.00)
November	0.72	0.66	0.68	0.60	0.87	0.83
	(0.00)	(0.00)	(0.05)***	(0.05)**	(-0.01)	(-0.01)
December	0.74	0.69	0.69	0.62	0.88	0.85
	(0.02)**	(0.02)**	(0.06)***	(0.06)***	(0.01)	(0.01)
REGIONAL VARIATION						
HH lives in	0.87	0.82	-	-	-	-
urban area	(0.15)***	(0.15)***	-	-	-	-
Dodoma	†	†	†	†	†	†
	(.)	(.)	(.)	(.)	(.)	(.)

Arusha	0.81	0.76	0.75	0.69	0.91	0.88
	(0.09)***	(0.10)***	(0.12)***	(0.14)***	(0.03)*	(0.04)*
Kilimanjaro	0.83	0.80	0.81	0.77	0.91	0.90
	(0.11)***	(0.14)***	(0.18)***	(0.21)***	(0.03)*	(0.06)***
Tanga	0.80	0.77	0.73	0.68	0.92	0.90
	(0.08)***	(0.10)***	(0.10)***	(0.12)***	(0.04)***	(0.06)***
Morogoro	0.75	0.71	0.67	0.62	0.90	0.87
	(0.03)	(0.05)**	(0.04)	(0.07)**	(0.03)	(0.03)*
Pwani	0.84	0.80	0.80	0.75	0.94	0.91
	(0.12)***	(0.14)***	(0.17)***	(0.19)***	(0.06)***	(0.07)***
Dar es Salaam	0.83	0.78	0.88	0.83	0.92	0.89
	(0.11)***	(0.12)***	(0.25)***	(0.27)***	(0.04)***	(0.05)***
Lindi	0.75	0.71	0.67	0.62	0.90	0.87
	(0.03)	(0.05)*	(0.04)	(0.06)*	(0.02)	(0.03)
Mtwara	0.72	0.67	0.65	0.60	0.88	0.84
	(0.00)	(0.01)	(0.02)	(0.04)	(0.00)	(0.00)
Ruvuma	0.69	0.64	0.60	0.55	0.87	0.83
	(-0.03)	(-0.02)	(-0.03)	(-0.01)	(-0.01)	(-0.01)
Iringa	0.71	0.65	0.65	0.59	0.87	0.81
	(-0.01)	(-0.01)	(0.02)	(0.03)	(-0.01)	(-0.02)
Mbeya	0.72	0.82	0.65	0.60	0.88	0.85
	(0.00)	(0.15)	(0.02)	(0.04)	(0.00)	(0.01)
Singida	0.75	0.69	0.64	0.58	0.91	0.86
	(0.03)	(0.02)	(0.01)	(0.02)	(0.03)*	(0.02)
Tabora	0.78	0.74	0.66	0.60	0.93	0.91
	(0.06)***	(0.08)***	(0.03)	(0.04)	(0.05)***	(0.07)***
Rukwa	0.74	0.69	0.67	0.62	0.88	0.85
	(0.02)	(0.03)	(0.04)*	(0.06)**	(0.00)	(0.01)
Kigoma	0.75	0.72	0.70	0.66	0.89	0.86
	(0.03)	(0.05)**	(0.07)**	(0.10)***	(0.01)	(0.02)
Shinyanga	0.71	0.64	0.62	0.56	0.88	0.83
	(-0.01)	(-0.02)	(-0.01)	(0.00)	(0.00)	(-0.01)
Kagera	0.71	0.66	0.65	0.60	0.86	0.82
	(-0.01)	(-0.01)	(0.02)	(0.04)	(-0.02)	(-0.02)
Mwanza	0.85	0.82	0.81	0.77	0.94	0.93
	(0.13)***	(0.16)***	(0.18)***	(0.21)***	(0.06)***	(0.09)***
Mara	0.89	0.86	0.84	0.79	0.96	0.95
	(0.17)***	(0.19)***	(0.21)***	(0.24)***	(0.08)***	(0.11)***

a) Commercialisation rate estimates based on the GLM estimates of the models reported in table 3.8.
b) The base category is defined as one with sample average for and assets. The base household is located in Dodoma and interviewed in January.
c) Deviations from the base occur singly.

* denotes that the dummy is significant at 5%; ** denotes that the dummy is significant at 1%
† denotes base category

As the theory predicts, the results show that commercialisation, i.e. how large a proportion of the household consumption is acquired through market transaction, is dependent on informal barriers for trade. Access to market information, here proxied by owning a radio, TV or a telephone, is statistically highly significant determinant for commercialisation, but its impact is limited to one percentage point on average and *ceteris paribus*. The statistical significance implies that access to information matters but relatively small magnitude suggests that owning a radio may not capture the important parts of market information. Social networks are likely to be more prominent sources of market information, and thus distance to the market is likely to reflect easiness of gaining market information directly from the traders. Distance to the market also reflects the effort of transport and time it takes to reach the market. As the distance increases, also the level of commercialisation falls rapidly. Already being two kilometres away from the market decreases the commercialisation by four percentage points in rural areas and two percentage points in urban areas, on average and *ceteris paribus.* As the distance increases, also the commercialisation rates drop in a non-linear fashion, and households located at least ten kilometres from the markets are likely to be 6-8 percentage points less commercialised than households within one kilometre from the nearest market, on average and *ceteris paribus*. The relative importance of even the first kilometres from the market highlights the fact that transport and travelling creates high transaction costs. As travelling on poor road network requires time and resources even smaller distances isolate households effectively from trade and trading information. The impact of distance is largely similar both in rural and urban areas.

Also the household demographics matter. As family size expands, household is more able to send one of its members to the market and thus, the larger the household, the more commercialised it is likely to be in urban areas. However, in the rural areas large family size means lower commercialisation as the household members are more likely to be able to produce most of the goods at home. Whether the household is headed by a man or a woman, on the other hand, has no statistical significance on household's commercialisation. Then again the education of the household head is statistically significant factor boosting household's commercialisation especially in urban areas. Better education is likely to increase the household's integration within the market as the head is better able to process market information. Primary education that is likely to proxy basic mathematical, reading, and writing skills facilitates household decision-making. Still, on the whole the size of this impact is rather modest varying from 0-3 percentage points on average and *ceteris paribus*.

Different indicators for agricultural activities decrease the estimated commercialisation throughout the line. If the main economic activity of the household head is farming, the household is likely to be 1-2 percentage points less commercialised on average and *ceteris paribus.* The modest effect is likely to reflect the commonness of farming: as the majority of the households are cultivating agricultural output, being a farmer looses its power to discriminate between more and less commercialised households. However, different types of agricultural activities and assets help to dig deeper into the determinants of commercialisation. Especially the size of land owned impacts commercialisation: landless households have very little possibilities to cultivate sufficient amount of food whereas households with more than ten acres of land are likely to cover their own demand for food and possibly even produce surplus. Furthermore, if the household owns cattle or sheep it is likely to be a relatively well-off farmer and thus it is more likely to produce sufficient amount of food at home and be less commercialised than a household who owns other types of assets. The magnitude of the assets as such, however, has little impact on commercialisation. Having access to main grid electricity, on the other hand, has a large positive effect on commercialisation: four percentage points in urban areas and up to thirteen percentage points in rural areas, respectively, on average and *ceteris paribus*. This may not be purely the result of electricity as such, but a proxy for household's location with respect to major roads and thus access to larger and more reliable markets. Households with electricity are also likely to be richer, and the electricity variable may then also indicate a wealth effect not already captured by the other variables.

In the full model the results show that commercialisation is significantly related to whether or not the household is selling either agricultural products or non-farm products and services. These impacts are large and well determined. Especially in rural areas a household is likely to be seven percentage points more commercialised if it engages in selling non-farm production, whereas in the pooled sample a household is likely to be five percentage points more commercialised. Selling agricultural production, on the other hand, decreases the commercialisation rates by five percentage points in the pooled sample, two in the rural sample and six in the urban sample, respectively, on average and *ceteris paribus*. If a household is selling agricultural production, it is likely to have surplus of food crops and be more than self-sufficient in food. As the consumption aggregate is constructed mainly on everyday consumption, a large part of it is food which in these households is produced at home decreasing the proportion of consumption acquired from the market. On the other hand, if the household sells non-farm products and services, it is more likely to have diversified its

income sources away from agriculture into other, possibly more profitable sectors and may not produce sufficient quantities of food but prefers buying it from the markets. The fact that the consumption aggregate constructed for the analysis is based on daily consumption items, i.e. mainly food, allows this conclusion to be tentatively drawn. Furthermore, one of the most influential determinants of commercialisation is having wage or business income as the main source of cash. These households are likely to sell their labour rather than use it for cultivating their own food. The impact is large[4] as the reference category is having agriculture as the main identified source of income. This highlights the importance and the potential of diversifying households' livelihood strategies beyond agriculture especially in the rural areas if further commercialisation and, indeed, consumption is to be achieved. The households in the comparison group, by definition, are likely to grow their own food and sell their surplus to the markets.

Seasonal variation of commercialisation raises an important point for discussion. The variation is much larger in rural areas, where the supply of goods is less reliable and the roads might be impassable during the rainy season. Still, the variation seems to be more related to the harvest season. The main harvest season between July and December boosts the goods available at the markets and also increases commercialisation rates substantially. On the other hand, early in the year during the ploughing and cultivation, the prices at the markets are likely to go up as the supply decreases and the households must rely on their own stocks. Unfortunately, the HBS did not collect price information which could be used to verify the seasonal price changes, and deriving unit prices from total values of the monthly consumption is likely to lead to estimation errors possibly greater than the seasonal variation in prices. Thus the speculation about prices presented here cannot be verified using the data at hand. In the urban areas, seasonal variation is much more limited as the markets are more reliable and stable, which also leads to more stable levels of commercialisation. The urban and rural commercialisation also seems to vary into opposite directions during the year: the harvest season boosts commercialisation in rural areas as the households gain cash income they can use to buy goods from the market. The urban commercialisation, on the other hand, is more stable.

Besides seasonal variation, also regional variation outperforms household characteristics as explanatory variables for commercialisation. This implies that region specific fixed effects are

[4] 11 percentage points in the pooled sample, 13 percentage points in the rural sample and 6 percentage points in the urban sample, on average and ceteris paribus.

larger determinants of commercialisation than the factors related to the individual households. Even after controlling for the households' access to the nearest market, the size of that market, reliability of the supply, and the links to other markets are crucial in households' decision to specialise. This further emphasises the importance of regional market structure, infrastructure and overall trade environment that are influencing the households' decision on market integration. Especially Arusha and Kilimanjaro with large markets near the Kenyan boarder, Tanga with its international harbour and good connections to both Kenya and Dar es Salaam, Dar es Salaam as the commercial capital, Mara and Mwanza next to the Lake Victoria with train connections and airports as well as direct access to Uganda, and finally Pwani in the immediate surroundings of Dar es Salaam are particularly favourable for trade.

In order to investigate whether the difference between the rural and urban level of commercialisation is mainly due to differences in endowments or treatment, these two effects are analysed individually. The most significant differences in the treatment are related to households' distance to the market, and variables related to farming, such as having farming as main activity, and owning land, cattle or sheep. Also seasons have different impact on commercialisation depending on whether the household is rural or urban. The source of the largest difference in the endowments, on the other hand, are the agricultural variables, as was to be expected. Not only are the rural areas more agriculturalised but farming has also different impact on commercialisation in rural areas compared to the urban areas. The second largest source of inequality comes from distance to the market: rural households are located further from the markets and each kilometre is also more harmful in terms of commercialisation than in the urban areas.

7 Conclusions

Commercialisation was found to be a significant determinant of household consumption in this study. Both the consumption side of commercialisation as well as engaging in trade as a seller are likely to lead to sizable increase in household's consumption. One percentage point increase in commercialisation leads to, on average and *ceteris paribus*, 0.3-0.5 percentage increase in the consumption in the pooled sample, 0.5-0.7 percentage increase in rural areas, and 0.1-0.4 percentage increase in urban areas, respectively. On the other hand, a household that engages in selling agricultural output consumes, on average and *ceteris paribus*, 11-13 percent more than a household that does not engage in trade in all samples. There is thus a clear case in favour of

enhancing further commercialisation in Tanzania. Besides commercialisation, also education and income sources had a particularly large impact on consumption, implying that commercialisation efforts should be a part of a larger development agenda where diversification of livelihoods, reliable market structures and social security networks are supported.

However, the means for achieving greater integration to the local markets and thus higher levels of commercialisation are not straightforward. Access to markets and market information were found to be significant determinants of commercialisation, and even shorter distances to the nearest market were found to isolate households and push them towards autarky. This finding encourages further investment in infrastructure. Other empirical evidence also backs up this recommendation: a recent evaluation of an infrastructure development programme in Tanzania shows large benefits from investing in rural roads that have previously been mostly impassable (Gajewski et al. 2002). Also household characteristics, such as education and inclination towards agricultural activities, were found to be important determinants of commercialisation, and thus further emphasis on primary and secondary education as well as improved opportunities for diversified livelihoods are ways to boost commercialisation and consumption. On the other hand, the size and significance of regional and seasonal fixed effects highlight the importance of structural barriers related to trade, that cannot be captured by household characteristics. Farmers' representatives and policy makers interviewed for this study confirm the notion of large costs for trade due to rigidities at the markets. For example, even though most of the farmers are able to access the closest trading point, the traders might not be there; farmers have poor knowledge on business practices; price fluctuation makes trade very unreliable; market information is poorly disseminated; credit is on short supply due to lack of collateral and unsure future profits from selling the harvest; traditional norms and government policy encourage food self-sufficiency; and poor processing and storage facilities are all examples of the structural constraints for further commercialisation facing Tanzanian households today. Thus, broad based development efforts are needed at all levels of the marketing chain to solve the issue of currently prevailing bottlenecks in trade in Tanzania.

At local village level the benefits from improved access to markets are conditional on appropriate policies being in place allowing and encouraging trade to develop. As the results of this study highlight there is substantial scope for improvement in the current trading environment to allow large scale commercialisation that would boost consumption levels throughout the country. As

Fafchamps et al. (2003) point out, trade is dependent on frictions in the whole marketing chain. Government officials can play an important role in ensuring that efficient marketing structures are in place and that legislation and actual praxis enables further reliance on the markets. Current efforts of the Tanzanian government to liberalise domestic trade, remove state monopolies and liberalise transportation markets are good examples of political measures to enhance possibilities to trade.

Due to the clear and large positive income effect of commercialisation found in this study, commercialisation efforts can be advocated as income enhancing in Tanzania. One possible and statistically significant tool to enhance commercialisation would be to improve the access to the markets and market information, but as discussed above, mere investment in infrastructure is not likely to lead to large changes in behaviour as there are still large structural constraints for households' commercialisation in place. The lack of transportation and poorly developed markets have made speculation possible even when the cost of transportation *per se* would be low. The fact remains that large regional fixed effects capture a lot of the variation in the model that cannot be otherwise captured by the household level variables. This emphasises the need for further research in order to understand the regional effects and structural barriers present in Tanzania restricting the households' commercialisation. Such research could also provide examples of especially favourable conditions for trading and models for best practice in enhancing commercialisation in order to improve the policies aimed at enhancing commercialisation and fight poverty.

References:

Airey, T. – Bryceson, D. – Howe, J. (1989) *Interim Evaluation of the Songea-Makambako Road*, Evaluation Department, Overseas Development Administration, London.

Ahmed, R. (1994) Investment in Rural Infrastructure: Critical for Commercialization in Bangladesh, in Von Braun, J. & Kennedy, E. (eds.) *Commercialization of Agriculture, Economic Development and Nutrition,* John Hopkins Press, Baltimore, MD.

Bardhan, P. & Udry, C. (1999) *Development Microeconomics*, Oxford University Press, Oxford.

Baum, C. – Shaffer, M. – Stillman, S. (2002) IVENDOG: STATA module to calculate Durbin-Wu-Hausman test after IVREG, STATA.

Bougheas, S. – Demetriades, P.O. – Morgenroth, E.L.W. (1999) Infrastructure, transport costs, and trade, *Journal of International Economics*, Vol. 47, Issue 1, pp. 169-189.

Datt, G. & Jolliffe, D. (1999) Determinants of Poverty in Egypt: 1997, *FCND Discussion Paper* No. 75.

Davidson, R. & MacKinnon, J.G. (1993) *Estimation and Inference in Econometrics*, Oxford University Press, New York.

De Janvry, A. – Fafchamps, M. – Sadoulet, E. (1991), Peasant Household Behaviour with Missing Markets: Some Paradoxes Explained, *Economic Journal* 101, pp. 1400-17.

Deaton, A. (1997) *The Analysis of Household Surveys, A Microeconometric Approach to Development Policy*, John Hopkins University Press, Baltimore.

Deaton, A. & Paxton, C. (1998) Economies of Scale, Household Size and the Demand for Food, *Journal of Political Economy*, Vol. 106, No. 7.

Durbin, J. (1954) Errors in variables, *Review of the International Statistical Institute*, Vol. 22, pp. 23-32.

Fafchamps, M. & Pender, J. (1997) Precautionary Saving, Credit Constraint, and Irreversible Investment: Theory and Evidence from Semi-Arid India, *Journal of Business and Economic Statistics*, Vol. 15, Issue 2, pp. 180-194.

Fafchamps, M. – Hill, R.V. – Kaudha, A. (2003) The Transmission of International Commodity Prices to Domestic Producers, The Centre of the Study of African Economies Working Paper Series, Paper No. 197.

Fan, S. (2003), Public Investment and Poverty Reduction – What have we Learnt from India and China?, Paper prepared for the ADBI conference, "Infrastructure Investment for Poverty Reduction: what do we know?" Tokyo, June 12-13, 2003.

Gajewski, G. – Wieland, R. – Johnson, R. – Mwasyoke, P. (2002) *Evaluation of Rural Roads Programme in Tanzania*, prepared for USAID/Tanzania, Final Report.

Gibson, J. & Rozelle, S. (2002), Poverty and Access to Infrastructure in Papua New Guinea, *Working Paper No. 02-008*, Department of Agricultural and Resource Economics, University of California Davis.

Gootaert, C. (2002) Socioeconomic Impact Assessment of Rural Roads: Methodology and Questionnaires, World Bank.

Hansen, L. (1982) Large sample properties of generalised methods of moments estimators, *Econometrica*, Vol. 50, No. 4, pp. 1029-1054.

Hausman, J. (1978) Specification tests in econometrics, *Econometrica*, Vol. 46, No. 6. pp. 1251-1271.

Hausman, J. & Leonard, G. (1997) Superstars in the National Basketball Association: Economic Value and Policy, *Journal of Labour Economics,* Vol. 15, No. 4, pp. 586-624.

Hayashi, F. (2000) *Econometrics*, Princeton University Press, New Jersey.

Huber, P. J. (1967) The behavior of maximum likelihood estimates under nonstandard conditions. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*. Berkeley, CA: University of California Press, vol. 1, 221–223.

Heltberg, R. & Tarp, F. (2002) Agricultural Supply Response and Poverty in Mozambique, *Food Policy*, Vol. 27(2), pp. 103-124.

Heywood, P.F. & Hide, R.L. (1994) Nutritional effects of export-crop production in Papua New Guinea: A review of the evidence, in *Food and Nutrition Bulletin*, Vol. 15, The United Nations University Press.

IFAD (2001) *Rural Poverty Report 2001: The Challenges of Ending Rural Poverty*, Oxford: Oxford University Press.

Key, N. – Sadoulet, E. – de Janvry, A. (2000) Transactions Costs and Agricultural Household Supply Response, *American Journal of Agricultural Economics*, Vol. 82, pp. 245-259.

Lambert, J. (1982). The Relationship between Cash Crop Production and Nutritional Status in Papua New Guinea. History of Agriculture. Working Paper 38. University of Papua New Guinea, Port Moresby. Cited in: J. Lambert, Effect of Urbanization and Western Foods, *Food Nutr. Bull*., 4(3). pp. 11-13.

Lanjouw, P. & Ravallion, M. (1995) Poverty and household size, *The Economic Journal*, Vol. 105, No. 433, pp. 1415-1434.

Larson, D. & Deininger, K. (2001) 'Crop Markets and Household Participation' in Reinikka, R. & Collier, P (eds) (2001) *Uganda's Recovery, The Role of Farms, Firms, and Government*, The World Bank, Washington D.C.

Lipton, M. & Ravallion, M. (1995) Poverty and policy, in *Handbook of Development Economics*, Vol. 3, edit. Jehrman, B. & Srinivasan, T.N., Noth-Holland Publishing Company, Amsterdam.

Longhurst, R. (1988) Cash crops, household food security, and nutrition, *IDS Bulletin*, Vol. 19, No. 2, pp. 28-36.

Lucas, K. – Rutachokozibwa, V. – Tagora, E. (1995) *The Njombe-Makete Road: An Impact Assessment of an ATAP Funded Road Improvement Project*, USAID, Dar es Salaam, Tanzania.

Maxwell, S. (2001) 'Agricultural Issues in Food Security' in Devereux, S. & Maxwell, S. (eds) (2001) *Food Security in Sub-Saharan Africa*, ITDB Publishing, London.

McKay, A. – Morrissey, O. – Vaillant, C. (1997), Trade Liberalisation and Agricultural Supply Response: Issues and Some Lessons, *European Journal of Development Research,* 9(2), 129.147.

Mtetewaunga, S.D. (1986) Social Implications of Land Reform in Tanzania, in Arntzen, J.W. – Ngcongco, L.D. – Turner, S.D. (edit.) *Land Policy and Agriculture in Eastern and Southern Africa*, United Nations University Press, Tokyo

Nakamura, A. & Nakamura, M. (1998) Model Specification and Endogeneity, *Journal of Econometrics*, Vol. 83, pp. 213-237.

NBS (2002a) *Household Budget Survey 2000/01*, Vol. V, National Bureau of Statistics, Dar es Salaam, Tanzania.

NBS (2002b) *Household Budget Survey 2000/01*, *Final Report*, National Bureau of Statistics, Dar es Salaam, Tanzania.

Ndumbaro, F.G.J. (1995), Rural Transportation in Tanzania: The Case of Mhukuru and Hanga Divisions in Songea Rural District, *Eastern & Southern Africa Geographical Journal*, Vol. 6, No. 1, pp. 37-48.

Omamo, S. (1998a) Transport Costs and Smallholder Cropping Choices: An Application to Sieye District, Kenya, *American Journal of Agricultural Economics*, Vol. 80, No. 1, pp. 116-123.

Omamo, S. (1998b) Farm-to-Market Transaction Costs and Specialisation in Small-Scale Agriculture: Explorations with a Non-separable Household Model, *The Journal of Development Studies*, Vol. 35, No. 2, pp. 152-163.

Papke, L.E. & Wooldridge, J.M. (1996) Econometric Methods for Fractional Response Variables with an Application to 401(k) Plan Participation Rates, *Journal of Applied Econometrics*, Vol. 11, pp. 619-632.

Park, J. & Davis, G. (2000) The Theory of Econometrics of Health Information in Cross-Sectional Nutrient Demand Analysis, Faculty Paper Series FP 01-02, Department of Agricultural Economics Texas A&M University College Station.

Pingali, P.L. & Rosegrant, M.W. (1995) Agricultural Commercialisation and Diversification: Processes and Policies, *Food Policy*, Vol. 20, No. 3, pp. 171-185.

Ravallion, M. (2001) The Mystery of the Vanishing Benefits: An Introduction to Impact Evaluation, *The World Bank Economic Review*, Vol. 15, No. 1, pp. 115-140.

Rogers, W. H. (1993) Regression standard errors in clustered samples. *Stata Technical Bulletin* 13: 19–23. Reprinted in *Stata Technical Bulletin Reprints*, vol. 3, 88–94.

Sargan, J.D. (1958) The estimation of economic relationships using instrumental variables, *Econometrica*, Vol. 26, No. 3, pp. 393-415.

Singh, I. – Squire, L. – Strauss, J. (edit.) (1986) *Agricultural Household Models: Extensions, Applications, and Policy*, The John Hopkins University Press, Baltimore.

Van De Walle, D. (2002a) Choosing Rural Road Investments to Help Reduce Poverty, *World Development* Vol. 30, No. 4, pp. 575-589.

Van De Walle, D. (2002b) Impact Evaluation of a Rural Road Rehabilitation Project, World Bank Discussion Paper.

Von Braun, J. & Kennedy, E. (1986), *Commercialisation of Subsistence Agriculture: Income and Nutrition Effects in Developing Countries,* Working Paper on Commercialisation of Agriculture and Nutrition 1, International Food Policy Research Institute, Washington, D.C.

Von Braun, J. (1995) Agricultural Commercialization: Impacts on Income and Nutrition and Implications for Policy, *Food Policy*, Vol. 20, No. 3, pp. 187-202.

Von Braun, J. & Kennedy, E. (eds.) (1994) *Commercialization of Agriculture, Economic Development and Nutrition,* John Hopkins Press, Baltimore, MD.

Williams, R. L. (2000) A note on robust variance estimation for cluster-correlated data. *Biometrics* 56: 645–646.

Wily, L.A. (2000) Land Tenure Reform and the Balance of Power in Eastern and Southern Africa, *Natural Resources Perspectives*, No. 58, Overseas Development Institute, London.

White, H. (1980) A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica* 48: 817–830.

Wooldridge, J. M. (2002) *Econometric Analysis of Cross Section and Panel Data.* The MIT Press, Cambridge, Massachusetts.

Wu, D.-M. (1973) Alternative tests of independence between stochastic regressors and disturbances, *Econometrica*, Vol. 41, No. 4, pp. 733-750.

Interviews:

Rosebud Kurwijila, Commissioner for Rural Economy and Agriculture, The Commission of African Union, 19.10.2004

Professor Simon M Mbilinyi, Member of Parliament, Chairman of the Tanzania Chamber of Minerals and Energy, 19.10.2004

Dr. Haji Mugishagwe, National Executive Chairman of Tanzania Eco-Development trust, 7.12.2004

Mr. E.R.K Mshiu, Chairman a& Managing Director, The Tanganyika Farmer's Association Ltd, Arusha, 18.10.2004

Appendix: COEFFICIENT ESTIMATES FOR HOUSEHOLD TOTAL CONSUMPTION PER ADULT EQUIVALENT FOR 28 DAYS (RURAL MODEL)

	Reduced OLS	Reduced IV	Full OLS	Full IV
Commercialisation	0.562***	-	0.488***	-
	(11.12)	-	(9.33)	-
Predicted values	-	0.658***	-	0.499**
	-	(3.37)	-	(2.37)
HOUSEHOLD DEMOGRAPHICS				
HH has only 1 member	†	†	†	†
	(.)	(.)	(.)	(.)
HH has 2-6 family members	-0.507***	-0.507***	-0.515***	-0.515***
	(15.59)	(15.10)	(15.87)	(15.46)
HH has 7 or more members	-0.866***	-0.868***	-0.885***	-0.888***
	(23.38)	(22.81)	(24.01)	(23.25)
HH head is male	-0.008	-0.006	-0.011	-0.011
	(0.39)	(0.31)	(0.59)	(0.56)
HH head age	0.000	-0.000	0.000	-0.000
	(0.11)	(0.27)	(0.01)	(0.38)
HH head sick o in last 4 weeks	0.007	0.005	0.010	0.009
	(0.41)	(0.28)	(0.59)	(0.50)
HH head has no education	†	†	†	†
	(.)	(.)	(.)	(.)
HH head has primary education	0.035	0.028	0.038*	0.033
	(1.58)	(1.27)	(1.75)	(1.52)
HH head has secondary educ.	0.036	0.033	0.042	0.041
	(1.03)	(0.94)	(1.20)	(1.17)
HH head has higher education	0.094**	0.091**	0.102***	0.099**
	(2.38)	(2.29)	(2.59)	(2.52)
MAIN ACTIVITY IS FARMING				
The main activity of the HH	-0.007	-0.005	-0.008	-0.007
head is farming/livestock	(0.37)	(0.23)	(0.43)	(0.35)
PRODUCTIVE ASSETS				
log of assets	-0.003	-0.002	-0.004	-0.002
	(0.56)	(0.34)	(0.70)	(0.46)
HH has no land	†	†	†	†
	(.)	(.)	(.)	(.)
HH owns max 2 acres of land	-0.015	-0.004	-0.010	-0.005
	(0.48)	(0.11)	(0.31)	(0.16)
HH owns 2-10 acres of land	0.036	0.050	0.045	0.051
	(1.07)	(1.31)	(1.29)	(1.42)
HH owns more than 10 acres	0.167***	0.174***	0.179***	0.181***
	(3.99)	(4.01)	(4.21)	(4.29)
HH has 1-10 cows	0.072***	0.072***	0.079***	0.078***
	(3.37)	(3.38)	(3.77)	(3.69)
HH has more than 10 cows	0.137***	0.149***	0.150***	0.159***
	(3.60)	(3.84)	(3.98)	(4.21)
HH has 1-10 sheep	0.010	0.005	0.001	-0.004
	(0.48)	(0.25)	(0.05)	(0.17)
HH has more than 10 sheep	0.023	0.023	0.013	0.014
	(0.71)	(0.72)	(0.42)	(0.44)
HH has access to electricity	0.308***	0.295***	0.284***	0.281***
	(6.99)	(5.82)	(6.64)	(6.00)
INCOME SOURCES				
HH sells agricultural output	-	-	0.113***	0.112***
	-	-	(5.27)	(5.26)
HH sells non-farm output	-	-	0.083***	0.078***

	-	-	(4.05)	(3.25)
MAIN SOURCES OF CASH				
HH's main source of cash is agriculture	† (.)	† (.)	† (.)	† (.)
HH's main source of cash is wage or business income	- -	- -	0.162*** (6.52)	0.167*** (4.74)
HH's main source of cash is remittances or other undefined	-	-	-0.024 (0.73)	-0.017 (0.53)
SEASONAL VARIATION				
January	† (.)	† (.)	† (.)	† (.)
February	0.025 (0.49)	0.023 (0.47)	0.017 (0.34)	0.014 (0.29)
March	0.063 (1.25)	0.073 (1.45)	0.051 (1.01)	0.059 (1.20)
April	0.077 (1.55)	0.075 (1.48)	0.062 (1.26)	0.060 (1.20)
May	-0.035 (0.63)	-0.044 (0.77)	-0.044 (0.81)	-0.048 (0.86)
June	-0.041 (0.79)	-0.039 (0.72)	-0.050 (0.99)	-0.045 (0.86)
July	0.045 (0.89)	0.038 (0.69)	0.032 (0.65)	0.030 (0.57)
August	0.041 (0.83)	0.035 (0.67)	0.027 (0.54)	0.024 (0.46)
September	0.026 (0.52)	0.020 (0.39)	0.011 (0.23)	0.011 (0.21)
October	-0.022 (0.43)	-0.033 (0.61)	-0.034 (0.69)	-0.039 (0.75)
November	0.046 (0.92)	0.044 (0.86)	0.034 (0.71)	0.036 (0.72)
December	0.065 (1.23)	0.054 (1.02)	0.065 (1.28)	0.062 (1.19)
REGIONAL VARIATION				
Dodoma	† (.)	† (.)	† (.)	† (.)
Arusha	-0.213*** (2.83)	-0.234*** (2.82)	-0.178** (2.31)	-0.189** (2.20)
Kilimanjaro	-0.161*** (2.81)	-0.175** (2.50)	-0.112* (1.89)	-0.109 (1.42)
Tanga	-0.068 (1.26)	-0.075 (1.25)	-0.037 (0.65)	-0.034 (0.54)
Morogoro	0.061 (0.86)	0.059 (0.80)	0.109 (1.55)	0.110 (1.50)
Pwani	-0.059 (0.75)	-0.076 (0.85)	-0.033 (0.43)	-0.033 (0.36)
Dar es Salaam	-0.096 (0.77)	-0.118 (0.88)	-0.086 (0.68)	-0.085 (0.63)
Lindi	-0.007 (0.07)	-0.011 (0.11)	0.026 (0.26)	0.029 (0.28)
Mtwara	-0.010 (0.09)	-0.008 (0.07)	0.035 (0.32)	0.042 (0.38)
Ruvuma	-0.104 (1.47)	-0.086 (1.23)	-0.067 (0.95)	-0.049 (0.70)
Iringa	0.033 (0.39)	0.013 (0.16)	0.062 (0.73)	0.045 (0.53)
Mbeya	0.120	0.115	0.150**	0.149**

	(1.64)	(1.60)	(2.00)	(1.99)
Singida	-0.259***	-0.236***	-0.223***	-0.201***
	(3.91)	(3.82)	(3.21)	(3.11)
Tabora	0.008	0.001	0.033	0.032
	(0.14)	(0.01)	(0.59)	(0.53)
Rukwa	-0.225***	-0.231***	-0.183**	-0.184**
	(2.93)	(2.96)	(2.40)	(2.36)
Kigoma	-0.173**	-0.178**	-0.144**	-0.141*
	(2.44)	(2.39)	(2.05)	(1.84)
Shinyanga	0.134*	0.139*	0.171**	0.175**
	(1.68)	(1.74)	(2.13)	(2.16)
Kagera	0.193***	0.194***	0.211***	0.215***
	(2.80)	(2.80)	(3.06)	(3.08)
Mwanza	-0.179**	-0.220***	-0.121	-0.143
	(2.31)	(2.58)	(1.56)	(1.61)
Mara	-0.398***	-0.419***	-0.358***	-0.358***
	(3.84)	(3.54)	(3.54)	(3.03)
Constant	9.442***	9.379***	9.362***	9.347***
	(89.47)	(62.88)	(90.66)	(66.23)
Observations	7580	7405	7580	7405
R-squared	0.21	0.25	0.22	0.21
RESET test	1.88	4.39***	2.15*	0.58
Prob >F	0.130	0.004	0.092	0.625
Hansen-Sargan	-	85.985	-	90.748*
Prob > Chi-sq	-	0.161	-	0.090
Endogeneity test	0.26	-	0.00	-
Prob > F	0.608	-	0.946	-

Robust t statistics in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%
† denotes base category

Appendix (continued): COEFFICIENT ESTIMATES FOR HOUSEHOLD TOTAL CONSUMPTION PER ADULT EQUIVALENT FOR 28 DAYS (URBAN MODEL)

	Reduced OLS	Reduced IV	Full OLS	Full IV
Commercialisation	0.418***	0.200	0.342***	0.084
	(8.13)	(0.93)	(6.63)	(0.38)
HOUSEHOLD DEMOGRAPHICS				
HH has only 1 member	†	†	†	†
	(.)	(.)	(.)	(.)
HH has 2-6 family members	-0.533***	-0.528***	-0.553***	-0.548***
	(26.95)	(25.87)	(28.55)	(27.68)
HH has 7 or more members	-0.909***	-0.904***	-0.943***	-0.938***
	(38.39)	(37.32)	(40.64)	(39.93)
HH head is male	-0.006	-0.008	-0.005	-0.007
	(0.47)	(0.60)	(0.42)	(0.55)
HH head age	-0.001***	-0.001***	-0.001***	-0.001***
	(3.42)	(3.47)	(3.36)	(3.39)
HH head sick in last 4 weeks	0.038***	0.036***	0.038***	0.037***
	(2.97)	(2.83)	(3.06)	(2.91)
HH head has no education	†	†	†	†
	(.)	(.)	(.)	(.)
HH head has primary education	0.006	0.009	0.008	0.011
	(0.38)	(0.50)	(0.50)	(0.64)
HH head has secondary educ.	0.017	0.022	0.019	0.023
	(0.74)	(0.89)	(0.80)	(0.98)
HH head has higher education	0.040	0.041	0.045*	0.045*
	(1.51)	(1.54)	(1.71)	(1.73)
MAIN ACTIVITY IS FARMING				
The main activity of the HH	0.025	0.020	0.029*	0.024
head is farming/livestock	(1.63)	(1.26)	(1.90)	(1.57)
PRODUCTIVE ASSETS				
log of assets	0.004	0.005	0.005	0.005
	(1.16)	(1.34)	(1.21)	(1.42)
HH has no land	†	†	†	†
	(.)	(.)	(.)	(.)
HH owns max 2 acres of land	-0.057***	-0.070***	-0.039**	-0.051**
	(3.38)	(3.18)	(2.37)	(2.55)
HH owns 2-10 acres of land	-0.027	-0.046*	-0.006	-0.021
	(1.35)	(1.66)	(0.31)	(0.86)
HH owns more than 10 acres	-0.002	-0.016	0.046	0.038
	(0.05)	(0.41)	(1.31)	(1.07)
HH has 1-10 cows	0.099***	0.096***	0.082***	0.081***
	(4.35)	(4.23)	(3.78)	(3.75)
HH has more than 10 cows	-0.089	-0.091	-0.080	-0.078
	(1.53)	(1.57)	(1.43)	(1.40)
HH has 1-10 sheep	0.061***	0.052***	0.049***	0.040**
	(3.39)	(2.60)	(2.74)	(2.00)
HH has more than 10 sheep	0.109***	0.102***	0.093***	0.086***
	(3.55)	(3.22)	(3.03)	(2.72)
HH has access to electricity	0.304***	0.313***	0.290***	0.297***
	(17.30)	(16.23)	(16.99)	(16.54)
INCOME SOURCES				
HH sells agricultural output	-	-	0.121***	0.108***
	-	-	(7.07)	(5.18)
HH sells non-farm output	-	-	0.137***	0.141***
	-	-	(10.32)	(10.15)

MAIN SOURCES OF INCOME				
HH's main source of cash is agriculture	† (.)	† (.)	† (.)	† (.)
HH's main source of cash is wage or business income	- -	- -	0.164*** (8.81)	0.181*** (7.83)
HH's main source of cash is remittances or other undefined	- -	- -	-0.021 (0.94)	-0.019 (0.84)
SEASONAL VARIATION				
January	† (.)	† (.)	† (.)	† (.)
February	0.025 (1.19)	0.022 (1.07)	0.029 (1.40)	0.026 (1.26)
March	-0.000 (0.01)	-0.003 (0.13)	0.004 (0.17)	0.000 (0.02)
April	-0.015 (0.64)	-0.019 (0.83)	-0.013 (0.59)	-0.018 (0.78)
May	-0.005 (0.18)	-0.008 (0.32)	-0.003 (0.13)	-0.008 (0.29)
June	-0.026 (1.00)	-0.029 (1.15)	-0.028 (1.09)	-0.031 (1.22)
July	0.029 (1.21)	0.024 (0.98)	0.022 (0.91)	0.017 (0.71)
August	0.002 (0.08)	-0.003 (0.11)	-0.006 (0.24)	-0.011 (0.45)
September	0.041* (1.68)	0.039 (1.60)	0.040* (1.65)	0.037 (1.55)
October	-0.012 (0.55)	-0.012 (0.55)	-0.022 (1.00)	-0.022 (1.01)
November	0.028 (1.27)	0.027 (1.24)	0.024 (1.10)	0.023 (1.07)
December	0.126*** (5.56)	0.127*** (5.59)	0.131*** (5.82)	0.133*** (5.85)
REGIONAL VARIATION				
Dodoma	† (.)	† (.)	† (.)	† (.)
Arusha	-0.123** (2.44)	-0.114** (2.19)	-0.094* (1.86)	-0.083 (1.59)
Kilimanjaro	-0.049 (0.90)	-0.038 (0.68)	-0.012 (0.22)	0.005 (0.09)
Tanga	0.025 (0.45)	0.038 (0.68)	0.071 (1.33)	0.088 (1.58)
Morogoro	0.108* (1.94)	0.116** (2.05)	0.139** (2.57)	0.149*** (2.71)
Pwani	-0.009 (0.17)	0.009 (0.16)	0.010 (0.18)	0.030 (0.53)
Dar es Salaam	0.093* (1.82)	0.104** (1.97)	0.131** (2.58)	0.142*** (2.73)
Lindi	0.062 (0.89)	0.070 (0.97)	0.085 (1.23)	0.095 (1.32)
Mtwara	-0.021 (0.31)	-0.016 (0.25)	-0.018 (0.27)	-0.013 (0.20)
Ruvuma	0.021 (0.34)	0.020 (0.32)	0.020 (0.33)	0.019 (0.30)
Iringa	0.244*** (3.82)	0.243*** (3.87)	0.254*** (4.22)	0.252*** (4.25)
Mbeya	0.119** (2.48)	0.120** (2.49)	0.176*** (3.59)	0.179*** (3.62)

Singida	-0.027	-0.021	-0.032	-0.027
	(0.52)	(0.39)	(0.60)	(0.50)
Tabora	0.251***	0.270***	0.291***	0.314***
	(3.75)	(3.91)	(4.64)	(4.77)
Rukwa	-0.031	-0.030	-0.007	-0.003
	(0.55)	(0.51)	(0.13)	(0.06)
Kigoma	0.252***	0.257***	0.253***	0.260***
	(4.39)	(4.43)	(4.43)	(4.49)
Shinyanga	0.267***	0.272***	0.278***	0.281***
	(5.40)	(5.37)	(5.83)	(5.77)
Kagera	0.256***	0.254***	0.267***	0.265***
	(4.87)	(4.82)	(5.01)	(4.96)
Mwanza	-0.008	0.010	0.040	0.062
	(0.13)	(0.15)	(0.66)	(0.97)
Mara	-0.225***	-0.199**	-0.145**	-0.114
	(3.11)	(2.54)	(2.07)	(1.48)
Constant	9.473***	9.653***	9.351***	9.551***
	(112.70)	(52.10)	(111.53)	(52.25)
Observations	14143	14125	14143	14125
R-squared	0.23	0.09	0.25	0.08
RESET test	2.13*	0.68	0.79	0.51
Prob > F	0.094	0.5671	0.500	0.674
Hansen-Sargan	-	62.598	-	59.048
Prob > chi-sq	-	0.254	-	0.365
Endogeneity test	1.09	-	1.43	-
Prob > F	0.296	-	0.232	-

Robust t statistics in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

† denotes base category