Adoption Pathways project discussion paper 1

November 2014

A 2008/2009 Social Accounting Matrix of Central and Southern Malawi

By

Sofie Skjeflo and Stein T. Holden

School of Economics and Business

Norwegian University of Life Sciences, Ås, Norway

Abstract

This paper describes a Social Accounting Matrix (SAM) that has been constructed to represent smallholder agricultural communities in six districts in the Central and Southern Regions of Malawi. The SAM has disaggregated rural households in groups based on region, sex of household head, and farm size. It captures the extent of access to farm input subsidies and informal markets for subsidized inputs. It shows the extent to which the different household groups are partially integrated into imperfect markets, mostly as net buyers of food and net sellers of labor. The SAM is intended as the basis for construction of a micro-meso Computable General Equilibrium model that can be used to analyze the direct and impacts of access to input subsidies and access to improved maize technology.

Jel codes: D57, D33, O33.

1. Introduction: A snapshot of a rural economy

A Social Accounting Matrix (SAM) is a square matrix that describes all payment flows in an economy in a given time period, where each cell in the matrix describes a payment from a column account to a row account. The accounts in the SAM represent production activities, commodities, factors and institutions, and the disaggregation of the accounts depends on the purpose of constructing the SAM. SAMs exist at the global, national, regional and village level, and are used in SAM multiplier analyses and as the basis for Computable General Equilibrium (CGE) models, where the disaggregation of the model follows the disaggregation of the SAM. The use of SAMs in the modeling of village economies in developing economies was pioneered by

Taylor and Adelman (1996), and the development of this SAM is in large part based on their work and others following their approach (Holden et al., 2005; Kuiper, 2005).

The purpose of this paper is to describe and discuss the development of a regional SAM of the rural economy consisting of farm households in six districts of central and southern regions of Malawi, which will be used as the basis for a CGE model of the same economy. Focusing on an economy consisting of farm households creates both opportunities and challenges. Narrowing down the focus from the national to the regional level allows a much more detailed disaggregation of the economy, with particular focus on the activities that constitute the livelihoods of the rural population we are interested in. On the other hand, rural economies in developing countries are characterized by transaction costs and imperfect markets that require special attention in modeling. This has to be taken into account when creating the SAM, as the CGE model is constrained by the database it is calibrated to. The treatment of imports, exports and the exchange rate is also different from a national SAM. The region does not have a separate currency, and the "Rest of the world" with which the regional economy trades refers to all areas and sectors not captured by the SAM, both in Malawi and abroad. The government account does not refer to the national government of Malawi, but rather the government activities that take place within the economy captured by the SAM.

This SAM represents households from the districts of Thyolo, Chiradzulu, Zomba and Machinga in southern Malawi and the districts of Lilongwe and Kasungu in central Malawi, with a population of just over 800 000 households in 2008 (National Statistical Office, 2010b). All households grow at least some maize for home consumption, as well as legumes and tubers. The most important cash crop is tobacco, which is more important in the central region than the southern region. Some households sell casual labor(*Ganyu* labor), and some engage in informal business activities. The households differ in their access to land and household labor, and a significant share is female-headed. Land scarcity is most severe in the southern districts, where population pressure is high and farm sizes decreasing. To capture household heterogeneity, households are disaggregated into six groups; female-headed households, male-headed land-poor households and male-headed land rich households in both regions. This disaggregation is further discussed in section 3 below. The same database and household categorization has been used by Holden (2013) to construct farm household models for simulation of household level impacts of access to improved maize technologies and input subsidies. While the household models typically capture average households for each household category, the SAM aggregates each household category based on the population size behind each household category and therefore illustrates the aggregate roles of each household category in the wider economy.

The SAM is highly aggregated for all non-farm activities, and can therefore not depict farm-non-farm sector linkages in detail. On the other hand, in line with Holden et al. (2005), all farm activities are disaggregated to the household group level in order to capture non-marketed consumption and transaction costs related to market participation. The disaggregation of time use into household chores, labor and leisure is also important for modeling the labor-leisure trade-off within the household as well as the particular characteristics of the labor market. The household-specific production activities may produce for household consumption (household-specific

commodity account) and for trade (aggregate commodity account). Some activities only produce output for sales, for instance the tobacco and the business activities, while some activities only produce output for own consumption, such as the household chores and leisure activities.

Transaction costs related to trade are captured through three accounts; the export transaction costs account, the import transaction costs account and the domestic transaction costs account. For each good that is traded domestically, exported or imported, an amount is paid to one of these accounts. The amount corresponds to the difference between the purchase price and the sales price of the good. For instance, when a household sells maize, the difference between the farm gate price and the market price of maize is the transaction cost. When a good is imported, the transaction cost is the difference between the lower international price and the higher price of the imported good when bought domestically. For exported goods, the transaction cost is the difference between the lower prices the exporter receives, and the higher price at which the good is sold abroad.

The Farm Input Subsidy Program (FISP) also requires special treatment in the SAM. Holden and Lunduka (2012) showed that there are in fact three different sources of fertilizers facing rural households. Through the FISP, some farmers could access subsidized fertilizer at less than 10 per cent of the commercial fertilizer price in 2008/2009. The intention of the program was to allocate two fertilizer coupons and one seed coupon to targeted households, entitling them to 100 kg of fertilizer at 800 MK per 50 kg bag, as well as seeds. Some households also received vouchers for tobacco fertilizers (Chibwana et al., 2010). However, a secondary illegal market for cheap fertilizer also developed, where fertilizer could be bought at 45 per cent of the commercial price (Holden and Lunduka, 2010). Finally, commercial fertilizer was also available at about 10 000 MK per 50 kg bag in the 2008/2009 season. This is captured in the SAM by differentiating between fertilizers accessed from the three sources. In the SAM, the government buys fertilizer and seeds at the commercial price, and then transfers some of it to the study area at the subsidized prices, while some of it is ”lost” through informal channels to traders. These traders transfer cheap fertilizer to the study area at the secondary market price, keeping the benefits from the transaction (the difference between the subsidized and secondary market price). The profits from the fertilizer subsidy to the beneficiary households are captured by the SubProfit account in the SAM.

2. Data and representative household categories

The detailed SAM is attached in a separate Excel file that we refer to as Table 1. It is divided in 168 accounts. The main data source of the SAM is household data from a survey of the 2008/2009 agricultural season, carried out in June 2009, the third round in a panel from the six above-mentioned districts in central and southern Malawi (Lunduka, 2009). Plot- and household level data from 373 households were used in the construction of the SAM, with the following allocation of respondents across the six household categories:

1: Female-headed south (17 per cent of sample)
2: Male land-poor south (24 per cent of sample)
3: Male land-rich south (17 per cent of sample)
4: Female-headed central (7 per cent of sample)

5: Male land-poor central (16 per cent of sample)
6: Male land-rich central (19 per cent of sample)

Household categorization is based on median owned land (inherited + endowed + purchased) available per household member. The area of land that is operated by the household is taken into account by using data on land rental activities. Households with less than median owned land per household member are defined as land poor. The median amount of owned land for the whole sample is at 0.155 hectares per household member. This categorization of households captures land scarcity when consumption needs and labor scarcity is taken into account. For land available for cultivation (after subtracting/adding rented out/in land) per household member, the median for the whole sample is 0.163 hectares. For either variable, there are more households in the land-poor category in the southern region than the central region. Table 2 shows summary statistics for each household group. In general, the female household heads are older, and have less schooling than the male household heads. Households are larger in the central region. As expected, the female-headed households have access to less male labor, and the land-poor households are more labor abundant. The median household owns just less than one hectare of land.

When classifying households for a SAM, the categories should be stable, in the sense that households are not expected to move in and out of the categories in response to shocks. Since land markets are relatively inflexible (at least in the short and medium run), and the labor endowment of the household group is also fairly stable in the medium run, a categorization based on land owned per household member seems reasonable. It is also important to distinguish between female-headed and male-headed households, since female-headed households may not face the same market access as male-headed households. For instance, Holden and Lunduka (2012) also showed that a smaller share of female-headed households received subsidized inputs through the FISP, despite being explicitly targeted in the program. Finally, the regional disaggregation of households allows us to take into account differential market and institutional characteristics of the central and southern regions, for instance differences in land scarcity and access to off-farm employment in the estate sector.

One problem when creating representative households is that these households appear to undertake many activities, but some of them only to a very small extent. For instance, the farm activities of each “representative household” based on an aggregated group appear to be much more diversified than for any real household. This could cause several problems, for instance underestimating vulnerability to adverse shocks by assuming too diverse income generating activities. Therefore, it is crucial to disaggregate households into as homogenous groups as possible, and carefully examine the data for outliers.

3. Separability and estimation of shadow prices

Agricultural household models describe households that are both producers and consumers of some of the same goods, for instance food crops, labor and land. In the agricultural household model with perfect markets, production decisions are made separately from consumption decisions, and the agricultural household can be modeled as a profit maximizing producer unit. Profits from farm production then enter the household’s decision making in the budget constraint of the utility

maximization problem. Non-separability in agricultural household models was introduced by Singh et al. (1986), who showed that when one or more markets for goods or factors that the household both demands and supplies fails, household production decisions are no longer independent from consumption decisions. In this case, the relevant price for decision-making in the household is not longer the exogenous market price, but the endogenous household-specific shadow price that equates household supply and demand of non-traded goods.

To enable more realistic modeling of the rural economy described here, it is vital to take into account market imperfections and the possible non-separability of consumption and production decisions. In particular, non-traded factors and commodities need to be valued at their shadow price in the SAM. Both the land and labor markets in the study area require special attention.

The land market in Malawi may be divided in the sales market and the rental market. Land sales are not very common but occur more often closer to towns such as in Lilongwe district. Land renting is more frequent although it has partly been suppressed by traditional norms (Holden et al. 2006). The SAM includes only land renting activity.

Only about 6 per cent of households report having bought at least one of their plots, and only 4 households report having sold a plot. The rental market is more active; 20 per cent of households report having rented in land while 11 per cent have rented out land. Almost 60 per cent of the rented out plots were rented out on a fixed rent contract, and 75 per cent of plots rented in were on fixed rent contract, with mean rental price per hectare of 2130 MK for land rented out and 1140 MK for land rented in.

Ganyu labor is various forms of short-term rural labor market opportunities such as temporary agricultural wage work on small farms or estates (Whitehead 2000). Such employment is particularly important for land-poor households that need supplementary income to meet their basic needs. Such employment is usually seasonally constrained and is typically offered in the peak season when farming households also need to work on their own farms. However, there are also non-agricultural *ganyu* employment opportunities that take place outside the peak agricultural season.

Our data indicate that the labor market is more active in the central region, more households are *Ganyu* workers during parts of the season, and they supply more labor than they demand. This may be because of the importance of the tobacco estate sector in Kasungu. In total, about 46 per cent of households report selling some *Ganyu* labor over the past year, while 35 per cent reports renting in some labor. As expected, land-rich households hire in more labor, while land-poor rent out labor. The *Ganyu* labor market is characterized by seasonality: there is limited access to off-farm labor in the peak season in agriculture, and limited access to off-farm employment in the low season.

3.1. Separability tests

As discussed above, when markets are well functioning, decisions regarding the

supply and demand of factors such as land and labor should be made independently of each other. This means that, controlling for the wage, factors that affect labor supply decisions should not affect labor demand (Benjamin, 1992). For instance, family size should not matter for the amount of labor used on the household's farm when markets are well functioning, rather all households apply the optimal labor input and any labor surplus or deficit is dealt with through the labor market. Optimal labor input only depends on the production technology and the (exogenous) wage. This result can be used to test whether the separability property holds for the households we observe in our data. Benjamin (1992) proposes the following test of separability:

$$H_0: Separability \rightarrow \frac{\partial F(L;A)}{\partial L} = w$$

where w is the exogenous market wage and $F(L;A)$ is the agricultural production function which depends on labor input L and a fixed amount of land A. The output price is normalized to one. The alternative hypothesis is non-separability, or

$$H_1: Non-separability \rightarrow \frac{\partial F(L;A)}{\partial L} \neq w$$

The implicit assumption is that the deviation from the equalization of the marginal product of labor to the exogenous wage depends on family size in the non-separable case. When labor supply and demand decisions are no longer separable, labor demand depends on an endogenous shadow wage, which depends on household composition. Benjamin (1992) presents three alternative sources of non-separability: a binding constraint on off-farm employment, rationing on the labor demand side and differing returns to labor on- and off-farm. Binswanger and Rosenzweig (1986) explain labor market imperfections in tropical agriculture by the seasonality in rain-fed agriculture and moral hazard related to hired labor. Feder (1985) models hired labor as an imperfect substitute to family labor due to moral hazard requiring monitoring of hired labor. Binding constraints on off-farm employment is likely to be the case in the low season for labor surplus households, while a binding constraint on access to labor could be the case for labor deficit households wishing to hire in labor during the high season. It may well be that there are differing returns to family and hired labor when used on the farm, or that there are better-paid off-farm opportunities, but either case is likely to be combined with labor shortages in the high season, and constrained access to off-farm employment in the low season. There may also be different skill requirements in off-farm employment causing low-skilled workers to be rationed out from such jobs.

An empirical problem that is also met by Benjamin (1992) is that we do not observe labor use separately for the peak season and the low season, but rather labor endowment available or total labor use over the year. Whether the test is able to detect non-separability then depends on the share of labor that is employed during the peak season. An additional problem is the potential endogeneity of demographic variables. For instance, unobservable or omitted variables such as land quality may affect both household size and labor demand on farm. Land quality (measured by plot fertility on a scale from 1 to 3 as stated by the farmer), whether the household received a coupon for cheap fertilizer, costs of purchased seeds and fertilizer and district dummies are included as an attempt to control for this potential source of endogeneity. The demographic variables included are household size and gender- and age composition of the households, as well as the education of the head of household. In line with

Benjamin (1992), we estimate the following empirical labor demand model:

$$\log L_h^D = \beta_0 + \beta_1 \log w_h + \beta_2 \log A_h + \beta_3 \sum_{i=1}^{G} \delta_i a_{i,h} + \varepsilon_h$$

where L^D are hours of labor used on the farm and G is the number of demographic variables. The regression results are shown in Table 3. The results show that….

3.2. Estimation of shadow prices

To estimate the shadow prices of land and household labor, we estimate the following aggregate agricultural production function, following Kuiper (2005), Jacoby (1993) and Skoufias (1994). This approach rests on the separability assumption that the production side decisions are not affected by the consumption side conditions for each household group.

$$\log Y_h = \alpha_0 + \alpha_1 \log Land_h + \alpha_2 \log Hhlab_h + \alpha_3 \log Ganyu_h + \alpha_4 \log Man_h + \alpha_5 \log T_h + \sum_i \beta_i X_{i,h} + \varepsilon_h$$

where for household h, Y_h is the value of agricultural output, valued at farm gate price, $Land_h$ is the amount of land cultivated, measured in hectares, $Hhlab_h$ are hours of household labor used in crop production, $Ganyu_h$ are hours of hired labor used in crop production, Man_h is kg of manure used in crop production, and T_h is the monetary value of traded inputs used: fertilizer, seed and pesticides. The tradable inputs are therefore treated in a very rough way and will need further disaggregation to handle the variation in fertilizer prices with access to input subsidies and the informal market for fertilizer. This requires a calibration approach rather than an estimation approach to achieve theoretical consistent results that satisfy exact aggregation, a requirement for the CGE-model. $X_{i,h}$ are controls for managerial ability (age and highest class completed by head of household), and district dummies. Summary statistics are shown in Table 4. The regression results are shown in Table 5.

All inputs have the expected positive sign and are significant in the production function, except for the manure variable, and the control variables have the expected signs. The estimated coefficients, $\hat{\alpha}_i$, are used to calculate shadow prices for land, labor and manure, using that the shadow price equals the value of the marginal product of the input. The shadow price is then calculated as

$$\hat{P}_{i,h} = \hat{\alpha}_i \frac{\hat{Y}_h}{N_{i,h}}$$

where $\hat{P}_{i,h}$ is the estimated shadow price for household group h for non-tradable i, $N_{i,h}$ is the amount of the non-tradable input used by the household group, and $\hat{Y}_h$ is the fitted output of household group h . The estimated shadow prices by household category are shown in Table 6.

We can expect the shadow value of family labor to be higher than the market wage if there are significant supervision costs, or if households have a preference for working on their own farm. In Malawi, however, the effect of limited access to off-farm employment and large transaction costs related to off-farm employment seem to dominate. In this case, we expect a lower shadow wage of family labor than off-farm labor. Another important aspect in the Malawi labor market is the seasonality in labor

demand, which we are not able to capture in the SAM which gives an average for the year. We do find lower annual average shadow wages for family labor than the *Ganyu* labor wage, with a mean shadow wage per hour of about 11 MK, versus a median *Ganyu* labor wage of about 30 MK/hour. This may also reflect that *Ganyu* labor is mostly employed at highly labor-intensive periods of crop production, while family labor is employed the entire year.

Access to land in Malawi is scarce and there is limited land rental and sales activities. Land is especially scarce in the southern region. We therefore expect the shadow value of land to be higher than the observed rental prices of land, and relatively higher shadow values in the South. The first hypothesis is supported (the mean shadow value of land is about 14 300 MK/ha while the cost of renting in land is about 1300 MK/ha). The farmers' median stated willingness to accept price to sell their plot is about 36 000 MK/ha.

Manure input was not significant in the production function regression, and the estimated shadow price of manure is not reliable. Less than half of households use manure, and manure use was not included in the SAM.

As expected, shadow prices are higher for households that are net renting in labor and land, and lower for households that are net suppliers of Ganyu labor.

3.3. Marketing margins for other goods

For crops that are both produced and consumed by the households, the difference between the farm gate price recorded in the survey and data on retail prices from Ministry of Agriculture and Food Security (2010) is used as an indicator of transaction costs. Jayne et al. (2010) compare retail prices from Ministry of Agriculture and Food Security (MoAFS) with farm gate prices obtained from focus groups in 2008 and 2009. The marketing margins they find are used as controls for the marketing margins used in this SAM. For goods that are both sold domestically and exported, we use the same transaction costs for exports as for domestic trade (implicitly assuming that the ROW retail price is the same as the domestic retail price). For goods that are both imported and produced domestically, we assume the same transaction costs for import as for domestic sales. For goods that are purely exported (except for tobacco for which we have price information at the farm gate and at the auction floors) or imported, we assume a 10 per cent transaction cost. The 10 per cent guess for goods with missing price data is a conservative guess if we compare to the marketing margins found in Jayne et al. (2010), where farm gate prices range from 61 to 96 per cent of retail prices.

Trade is assumed to be carried out by external traders that are not part of the household sample. The transaction costs are absorbed in the "commodity" CTrad, which is an imported transaction service commodity.

One exception is the transaction costs related to *Ganyu* employment by households. Here, we assume that selling *Ganyu* labor requires a transaction cost of x hours of household labor per hour of *Ganyu* labor supplied. *Ganyu* employment is included as a separate activity in the SAM, with payments to the household specific *Ganyu* commodity as well as the household specific time factor to take into account transaction costs that explain the difference between the market *Ganyu* wage and the

household shadow wage.

3.4. Profits from fertilizer subsidy and land rental

The observed prices on subsidized fertilizer and rented land imply that there is a discrepancy between the shadow value of these commodities at the household level and their market price. Households have constrained access to these inputs and constrained profit-maximization can be used to calibrate production functions where the gap between the commercial price and the subsidized price and the informal market price lead to profits for the household groups that benefit from access to such inputs at reduced prices. Holden et al. (1998) have shown how shadow prices may be calibrated with a 2-level CES production function, an approach we will used to calibrate the CGE-model linked to the SAM as well. Since the main purpose of constructing this SAM is to use it for development of a micro-meso CGE model it is important that the SAM is constructed and calibrated in a way that is consistent with the structure and theoretical formulation of agent behavior in the CGE model. This also complicates the SAM-construction.

For households that rent in land, the difference between the low rental price and the high shadow value of this land is returned as profits. For households that receive subsidized fertilizer, the difference between the low subsidized price and the commercial price of the fertilizer is returned to the household as profits and captures the income effect of access to subsidized inputs.

4. Elements of the SAM

4.1. Production activities and commodity markets

Harvest

Harvest from the crop activities is summarized in Table 7. There seems to be much less production of local maize varieties in the central region than the southern region, and in general the maize harvest is slightly larger in the central region. On average, between 55 and 64 per cent of households in the southern region cultivate local maize varieties while between 30 and 39 per cent cultivate local varieties in the central region. For improved maize the situation is reversed. Households in the Central region also cultivate more legumes than households in the South, where legumes are more often cultivated mixed or intercropped with maize. Female-headed households in both regions produce less tobacco than the male-headed households.

For simplicity, we assume that local maize activities in the southern region give the reported maize harvest, plus a legume harvest equal to 10 per cent of the maize harvest. In constructing the SAM, the recorded harvest is lower than the recorded gross output (value at input prices) for the female-headed and the large-scale households in the south, suggesting that green harvest and legume harvest from intercropping are under-reported. From the plot level data we see that intercropping and mixed cropping with both local and improved maize is much more common in the southern region, where almost 50 per cent of households have at least one plot that is intercropped local and improved maize. The maize production activities of the southern households thus have several outputs; legume for own consumption, maize for own consumption and maize for sale (although none of the household groups in

the south are net sellers of maize).

Inputs

The inputs used in production are summarized in Table 8. We have data on labor days by household members (number of members involved in activity times days spent on activity), for land preparation, planting, fertilizer application, weeding, harvesting and manure application, hired labor (number of workers times days hired). We assume 5 hours per man-day, and recalculate the labor input as hours. About 35 per cent of households used one or more days of *Ganyu* labor on their farm, giving zero amount for hired (*Ganyu*) labor for five out of six household groups in Table 8.

The land-rich households use less labor per hectare than the land-poor, which is in line with the hypothesis of seasonal constraints in the labor market, as well as imperfect land and labor markets. Land-rich households also use less fertilizer per hectare than the land-poor and female-headed households within the same region, and households in the south use higher levels of fertilizer per ha. This indicates that land scarcity leads to land use intensification.

Fertilizer and seeds used in maize production for each household group are only disaggregated according to whether the inputs are subsidized, purchased with coupons (likely to include some coupons achieved through informal systems), or purchased elsewhere (likely to include cheap inputs purchased through informal channels). The share of inputs obtained through purchased coupons is available through the NCA data at the household level, and Holden and Lunduka (2010) impute the amount of fertilizer obtained through the informal market for cheap fertilizer. The numbers they report imply that across all households, about 55 per cent of all fertilizer is subsidized and obtained through the official channel, 25 per cent is obtained through secondary markets (both through the purchase of coupons and the purchase of cheap fertilizer) and at a price between the full subsidy price and the commercial price, and 20 per cent is obtained through regular commercial markets. About 65 per cent of households in the NCA survey report (Holden and Lunduka 2010a) receiving input coupons in the 2008/2009 season. In the NOMA data, collected from the same households the following year , about 76 per cent of households report using subsidized fertilizer on at least one plot (Holden and Lunduka 2010b). The discrepancy between these numbers could be due to underreporting in the NCA survey, or the NOMA data could also include the use of subsidized fertilizer obtained through secondary markets, and some discrepancy in the samples due to attrition. Households in the southern region obtain a larger share of their fertilizer through the subsidy program, and a smaller share from commercial markets. Only small quantities of seed seem to have been obtained from the subsidy program; of those who reported getting seeds through the subsidy program (about 40 percent of sampled households), the average quantity was 2.8 kg. On average, about 16 percent of improved maize seeds were obtained through the subsidy program. The two data sources were consolidated for each household to get as reliable information as possible. Chibwana et al. (2010) report that about 4 percent of households in Kasungu and Machinga received vouchers for tobacco fertilizer, while in our data, 14 percent of households use subsidized tobacco fertilizer on at least one of their plots.

Since there is no information on maize or other crops retained as seeds, it is assumed that all improved maize seeds are purchased while all local maize seeds are from own

production.

4.2. Household incomes and expenditures

4.2.1. Income sources and consumption of own output

Consumption of own output

Table 9 shows summary statistics for household group consumption of own produce. We assume that food consumption requires some processing of foods, which takes place in the household chores activity. This household specific activity uses all home-consumed crops and all purchased foods as input, as well as labor, to produce a household group specific commodity (e.g. meals) that is consumed by the household. This commodity is entirely for home consumption.

Household time use, endowments and income sources

The available time use data from 2008/2009 is summarized in Table 10. The variables *peaklabor* and *leanlabor* are measured in mean hours per day spent in field per household member. The rest of the variables are measured as share of time during a year allocated to each activity. Since it is not clear whether this data is for the household head, and how representative it is for the time use of the whole household, it should be complemented with other data sources. We also have plot-level data on the number of man-days spent on each plot, which we have used as labor input use at the crop level below. For leisure and household chore activities, we use the detailed data on time use in Zambian households from Holden (1991). This has also been used in the household modeling for the same categories of households (Holden 2013).

Information on *Ganyu* employment and formal wage employment is summarized in Table 11. On average, 14 per cent of the households supply some formal labor, and 46 per cent supply some *Ganyu* labor. The land-poor households in the central region have the largest share supplying *Ganyu* labor; 70 per cent, and in general, the households in this region supply more *Ganyu* than in the southern region. The *Ganyu* wage is higher in the central region. Land-poor households supply more labor, both formal and informal, than the land-rich in both regions.

To capture the trade-off between labor and leisure at the household group level, a separate time activity is included in the SAM. This household group-specific activity uses the household group-specific time factor as input, and the output is allocated to household group *Ganyu* supply, and the leisure commodity, which is household group-specific. The leisure commodity is entirely consumed by the corresponding household group.

Summing up labor demand from all production activities and comparing with labor supply (both on- and off-farm), shows that the economy described by the SAM is a net exporter of labor, perhaps to the estate sector that is not covered by the SAM. The survey and data collection took place before it was known that the data should be used for SAM construction and CGE modeling. This implies that the data on sinks and sources for various transactions are incomplete. We therefore have to apply common sense and make some assumptions in the SAM construction and modeling work.

Household business activities

Household business activites include household enterprises and formal labor activities. Investments, average monthly profits and duration of operation (months per year) are summarized in Table 12. Unfortunately, this data is only collected for enterprises that were active over the past month at the time of the survey. Most business owners report acquiring start-up capital from own savings from agriculture (33 per cent), while the rest obtained capital from loans or gifts from family or friends (19 per cent), proceeds from other businesses and savings from *Ganyu* (10 per cent each), while the rest inherited, used other savings or borrowed from money lenders or banks (very few). We see that a smaller share of female-headed than male-headed households engage in business activities; 19 and 23 per cent in the two regions compared to 33 percent in the entire sample. Investments seem to be higher for land-rich households, while there is no clear pattern for profits (this data shows a lot of variation).

In the SAM, we have assumed that the household groups only used their own labor as input in the business activity. Labor input is calculated by assuming that all time spent on ”other non-agricultural activities” by the households engaged in business activities, is used as input into this activity for the duration of the business enterprise. The operating costs are accounted for as expenditure on other imported goods. Some inputs could in fact be locally available, for instance for food processing businesses, but due to lack of data we cannot specify this. Investments are paid from the household specific capital accounts.

Household non-crop activities

These non-crop activities include forest and livestock activities. We assume that these activities do not take place on household plots, but on communal land and forest areas not owned by the household implying weak linkages between crop and livestock production activities. Holden and Lunduka (2012) found no connection between livestock production and application of organic manure on crops. The inputs into these activities are therefore mainly household labor and capital (for the livestock activity). We assume that all value added from the forest activity is to labor, while 25 per cent of return from the livestock activity is to labor and the rest to capital. In the survey, households are specifically asked about home consumption of livestock products and forest harvest, and the value of this consumption. We assume that the marketed surplus from these activities is exported, so the households are either net sellers or self-sufficient in these products.

Household assets and productive capital

Since we do not have data on capital use in productive activities, we use the value added to capital (as shares of gross output) from the 2004 Malawi SAM from IFPRI (Thurlow et al., 2008) to impute the value added to capital in this SAM. We multiply the shares from the 2004 SAM with output valued at farm gate prices to arrive at the value added in the new SAM (We use output valued at farm gate prices to avoid amplifying errors in the other input-output entries). We assume that the shares vary across crop activities, but not across household types.

Household transfers from ROW/ROM

The survey data contains information on the value of gifts such as cash, food or other in-kind gifts, received from people outside the household. Unfortunately, there is no information on the source of the gifts. As a simplification, we assume that all transfers are received from outside the SAM as a transfer from ROW (Rest of the World, including Rest of Malawi (ROM) (e.g. remittances from migrated households).

To summarize, household incomes are derived from:

- Total value added to land from crop activities
- Total value added to capital from crop, non-crop and business activities
- Total value of their time endowment, valued at the shadow wage of labor
- Transfers from other households, assumed outside SAM area

4.2.2. Household expenditures

Expenditure data for food and non-food items with one month recall is recorded in July, which is during/after the main harvest. It is therefore likely that the data we have underestimates food purchases, and overestimates purchases of other goods and luxury goods if we multiply these by twelve to estimate annual expenditures. Table 13 shows summary statistics of household expenditures for the following commodity categories:

- Aggregate traded maize commodity (also produced by household)
- Aggregate traded legume commodity (also produced by household)
- Aggregate traded tuber commodity (also produced by household)
- Aggregate traded other crops commodity (also produced by household)
- Aggregate non-food commodity (not produced by household, imported)
- Aggregate food commodity (not produced by household, imported)

In order to adjust the expenditure data to reflect annual consumption, we use information from the NCA survey where households were asked about total annual maize consumption and production (Holden and Lunduka 2010a). We adjust food purchases upward by the same factor as the ratio of NCA maize expenditures to NOMA maize expenditures (Holden and Lunduka 2010b), except for household groups 4 and 6 for which these expenditure data were similar. Summary statistics in Table 14 show the distribution of net sellers and net buyers of maize across household groups, and this information is used to adjust maize expenditures. We see that the only group of households that on average were net sellers in the 2008/2009 season was the land-rich household group in the central region. The variable Net sell/buy 2009 takes the value 1 for net sellers, 2 for self-sufficient households and 3 for net buyers. As expected, there is a larger share of net buyers among female-headed households and land-poor households, and the land-rich in the central region have a mean value below 2, indicating a large share of net sellers.

By combining the information on crop sales at the household level, we find that about half of the surveyed households sell at least some of their output in the market, as shown by the summary statistics of the dummy variable *Seller*, in Table 14. This is slightly lower than the estimate in Wood and Mayer (2006) at 54 per cent at the national level, based on data from IHS2. They also find that the share that sells some output increases with income and with land holdings. Again, we see that a larger

share of the land-poor households in the south sell their crops than the land-rich, which is surprising. For the Central Region, the results are as expected, with the share of sellers increasing with land available.

We also combine information from IHS2 on rural household food expenditures to impute other food expenditures. The module on food consumption contains information on consumption of food over the past week, and since the survey is undertaken over a whole year to avoid seasonality problems, this data is likely to be more reliable than our expenditure data. Sampling strategy and survey are described in National Statistical Office (2005). We use the data on rural households from central and southern regions to make the sample comparable to the sample from 2008/2009. We then calculate food purchases as a share of total consumption for each of the crop categories in the SAM (i.e. for the crop groups that are both produced and consumed by the household). The results are shown in Table 15, and show that on average, households purchase about 40 per cent of the maize they consume. We used the expenditure shares of food categories in Ecker and Qaim (2011), which is based on the IHS2 survey data, to control our data. Unfortunately, these data are not disaggregated across household types, so we implicitly assume the same expenditure shares across households. These are also shares of all expenditures, including imputed value of home consumption and in-kind transfers. These shares are shown in Table 16.

4.3. Government, Savings, Investments and the Rest of the world (ROW)

There is no available data on local government activities (tax collection, transfers, etc) except for the input subsidies. All other government activities have to be based on outside data sources. We used tax rates from the 2004 SAM where the rural households in the 4th and 5th income quintiles pay direct taxes to the government, at rates of 0.3 per cent and 4.2 per cent, respectively. Since our SAM includes much more informal income, we apply the lowest tax rate to the land-rich households in both regions. We used net purchases of assets and livestock to impute savings rates for households.

We assumed a withholding of tax on tobacco auction floors of seven per cent, but smallholders in clubs or farmer associations were exempted. From 2010, this was reduced to 3 per cent, but with no exemptions (Chirwa, 2011).

5. Balancing and aggregation

5.1. Balancing the SAM

Balancing takes place both before and after aggregation. When organizing and cleaning the raw data, the data is thoroughly checked for data entry errors that may have created outliers. The cleaned data is then disaggregated to the representative household group level, and used to create the separate tables that in the end are combined as the complete SAM. Each of these sub-tables is balanced before aggregation. For instance, the input-output table is balanced to make gross activity outputs from crop activities match the value of reported harvest. We assume that the harvest data is more reliable than the input data, because the input data is recorded at the plot level rather than the crop level. The input data therefore does not take into

account inter-cropping and mixed cropping systems, but rather assumes that all inputs used on a plot are used as input into the primary crop planted. This creates imbalances that are particularly evident for the inputs in the maize cropping activity. All adjustments are made so that the relative positions of the household groups are maintained. Trade positions are netted out within each household group. since SAM cannot handle seasonality such that buying and selling activity over time for households also is netted out for the year. We therefore eliminate within-group trade by subtracting the quantity sold from the quantity purchased, revealing the household group as a net seller, net buyer or self sufficient of each good.

Adjustments in input-output table

Since the labor input data seems to be the least reliable, and plot size is measured by GPS and should therefore be relatively more reliable, we adjust labor input by assuming fixed or typical standardized amounts of labor per hectare for each crop type. We use the same labor input intensities as in the Malawi household models (Holden 2013). We cross-checked our survey data with the report by Simtowe et al. (2010), based on data from the 2006/7 crop season, and data from the household model that used our household plot panel survey data from 2005/06, 2006/07, and 2008/09 (Holden 2013).*Balancing incomes and expenditures*

Because of the poor quality of the consumption expenditure data, we choose to adjust this expenditure data rather than the income data to balance incomes and expenditures at the household group level (again because of the poor data quality, it is not feasible to balance incomes and expenditures at the household level before creating the average representative household for each group). After all other adjustments are made expenditures are still high compared to incomes for all household groups, probably indicating some hidden incomes. Adjusting expenditure involves adjusting cash expenditures, since subsistence expenditures are by definition balanced against income from farm activities that are consumed on-farm. Since the expenditure data was collected just after harvest, a downward adjustment of non-food expenditure seems reasonable, and is here used to balance the data rather than an ad hoc downward adjustment before balancing.

5.2. Population data, aggregation and representativeness

The report by National Statistical Office (2010b) contains data on the distribution of households and household size at the district level. From this data, we use the shares of households in each category (1-3 in each of the four southern districts, and 4-6 in the two central districts) to allocate the households in each of our sampled districts to a household category. This is based on the assumption that each of our household group samples is representative at the group in the districts in the two regions and therefore by weighting the groups with the population we have a model that is representative of these six districts. We also assume that these districts are fairly representive of the broader regions such that it helps us to assess the situation of typical smallholder households in the Central and Southern Regions of Malawi.

In order to assess the representativeness of the SAM, we compared the area planted of each crop category with area planted by crop from the 2006/2007 National Census of Agriculture and Livestock (NACAL) (National Statistical Office, 2010a). The results are shown in Table 17. It appears that in particular our sample from Southern Region

has higher production area data than was found in the NACAL survey. This is the case both for maize production area and tobacco production area. We need to further scrutinize and examine this.

6. Conclusion

We have described how we have constructed a Social Accounting Matrix (SAM) for smallholder agriculture in six districts in the Central and Southern Regions of Malawi. The SAM is disaggregated in household groups by region, sex of household head and farm size. We find that all but one of six household groups are net buyers of food with maize being the main staple crop. The SAM is particularly detailed in its disaggregation of production activities, including access to subsidized inputs through the formal targeting system as well as through the informal market. The SAM can be used to assess the distributional implications of the input subsidy program. As a basis for a micro-meso Computable General Equilibrium (CGE) model it can be used to assess the direct and indirect effects of access to input subsidies and improved maize technologies. The CGE model may be used to capture in particular the indirect price effects such as the effects through changes in endogenous maize prices and wage rates.

References

Benjamin, D. (1992) Household composition, labor markets, and labor demand: testing for separation in agricultural household models. *Econometrica*, 287-322.

Binswanger, H. P., & Rosenzweig, M. R. (1986). Behavioural and material determinants of production relations in agriculture. *The Journal of Development Studies* 22(3), 503-539.

Chibwana, C., Fisher, M., Jumbe, C., Masters, W., Shively, G. (2010) Measuring the Impacts of Malawi's Farm Input Subsidy Program. Available at SSRN 1860867.

Chirwa, E.W., (2011) Analysis of the Tobacco Industry in Malawi, United Nations Conference on Trade and Development.

Ecker, O., Qaim, M. (2011) Analyzing nutritional impacts of policies: an empirical study for Malawi. *World Development* 39, 412-428.

Feder, G. (1985). The relation between farm size and farm productivity: The role of family labor, supervision and credit constraints. *Journal of Development Economics* 18(2), 297-313.

Holden, S. (1991) Peasants and sustainable development–the chitemene region of Zambia–theory, evidence and models. Unpublished PhD thesis. Department of Economics and Social Sciences, Agricultural University of Norway, As.

Holden, S. (2013). Input subsidies and demand for improved maize: Relative prices and household heterogeneity matter! CLTS Working Paper No. 6/2013. Centre for Land Tenure Studies, Norwegian University of Life Sciences, Ås, Norway.

Holden, S. T., Kaarhus, R. and Lunduka, R. (2006). Land Policy Reform: The Role of Land Markets and Women's Land Rights in Malawi. Noragric Report No. 36, Noragric, Norwegian University of Life Sciences (UMB), Ås.

Holden, S., Lofgren, H., Shiferaw, B., (2005) Economic reforms and soil degradation in the ethiopian highlands: a micro CGE model with transaction costs, Paper presented at International Conference on Policy Modeling, Istanbul, June.

Holden, S., Lunduka, R. (2010) Impacts of the fertilizer subsidy program in Malawi: Targeting, household perceptions and preferences. Report to NORAD. Department of Economics and Resource Management, Norwegian University of Life Sciences.

Holden, S., Lunduka, R. (2012) Do fertilizer subsidies crowd out organic manures? The case of Malawi. *Agricultural Economics* 43, 301-312.

Holden, S. T. and Lunduka, R. (2013). Who Benefit from Malawi's Input Subsidy Program? *Forum For Development Studies* 40(1), 1-25.

Holden, S. T., Taylor, J. E., and Hampton, S. (1998). Structural Adjustment and Market Imperfections: A Stylized Village Economy-Wide Model with Nonseparable Farm Households. *Environment and Development Economics* 4: 69-87.

Jacoby, H.G. (1993) Shadow wages and peasant family labour supply: an econometric application to the Peruvian Sierra. The Review of Economic Studies 60, 903-921.

Jayne, T.S., Mangisoni, J., Sitko, N., Ricker-Gilbert, J. (2010) Malawi's Maize Marketing System. Report commissioned by the World Bank and Government of Malawi/Ministry of Agriculture, Lilongwe.

Kuiper, M.H., (2005) Village Modeling - A Chinese Recipe for Blending General Equilibrium and Household Modeling. Wageningen University.

Lunduka, R.W., (2009) Land rental markets, investment and productivity under customary land tenure systems in Malawi, Department of Economics and Resource Management. Norwegian University of Life Sciences, Ås, Norway.

Ministry of Agriculture and Food Security, (2010) Malawi Market Prices (Nominal Retail Market Prices), January to December 2009, in: Ministry of Agriculture and Food Security, G.o.M. (Ed.), Malawi.

National Statistical Office (2005) Integrated Household Survey 2004-2005, Socio-Economic Characteristics. 1.

National Statistical Office, (2010a) National Census of Agriculture and Livestock 2006/7, Main Report. National Statistical Office, Zomba, Malawi.

National Statistical Office, (2010b) Statistical Yearbook. National Statistical Office, Zomba, Malawi.

Simtowe, F., Asfaw, S., Shiferaw, B., Siambi, M., Monyo, E., Muricho, G., Abate, T., Silim, S., Ganga Rao, N., Madzonga, O. (2010) Socioeconomic Assessment of

Pigeonpea and Groundnut Production Conditions–Farmer Technology Choice, Market Linkages, Institutions and Poverty in Rural Malawi.

Singh, I., Squire, L., Strauss, J. (1986) Agricultural household models: Extensions, applications, and policy. Johns Hopkins University Press Baltimore.

Skoufias, E. (1994) Using shadow wages to estimate labor supply of agricultural households. American Journal of Agricultural Economics 76, 215-227.

Taylor, J.E., Adelman, I. (1996) Village economies: The design, estimation, and use of villagewide economic models. Cambridge University Press.

Thurlow, J., Diao, X., McCool, C., (2008) 2004 National Social Accounting Matrix (SAM) for Malawi, in: Institute, I.F.P.R. (Ed.), Washington D.C.

Whiteside, M. (2000). Ganyu Labour in Malawi and its implications for Livelihood Security Interventions: an analysis of recent literature and implications for poverty alleviation. London: Overseas Development Institute.

Wood, A., Mayer, J., (2006) Malawi Poverty and Vulnerability Assessment: Investing in our Future. Washington, DC: World Bank.