

TI 2016-054/V
Tinbergen Institute Discussion Paper



How does Market Access for Smallholders affect Export Supply? The Case of Tobacco Marketing in Malawi

Revision: August 2018

Wouter Zant¹

¹ VU Amsterdam, The Netherlands

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

Contact: discussionpapers@tinbergen.nl

More TI discussion papers can be downloaded at <http://www.tinbergen.nl>

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 598 4580

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900

How does Market Access for Smallholders affect Export Supply?

The Case of Tobacco Marketing in Malawi

by Wouter Zant

Abstract

We measure the impact of a reduction in transaction costs on crop area and production decisions of smallholder tobacco growers in Malawi. For identification we exploit the introduction of an additional tobacco auction floor. Estimations are based on annual data by Extension Planning Area. A 10% reduction in distance to auction floor is shown to increase crop area and production around 4% and 10% respectively. Supply response weakens beyond a distance to auction floor of 60km and runs along the intensive margin: existing tobacco growers improve productivity of cultivation. Impacts are robust for non-random placement of auction floor and several other threats.

JEL code: D23, O13, O55, Q11, Q13

Key words: transaction costs, market access, subsistence, food & cash crops, Malawi, Africa

* Wouter Zant is associate professor at the Vrije Universiteit and research fellow of the Tinbergen Institute, both Amsterdam, the Netherlands; mailing address: Wouter Zant, Vrije Universiteit, De Boelelaan 1105, room 10A-79, 1081 HV Amsterdam, The Netherlands; email: wouter.zant@vu.nl; tel: +31 20 598 9592.

Introduction

Smallholders in developing countries can choose to produce food crops for home consumption or cash crops for the market¹. High production costs, high transaction costs, and high risks of output and input prices often make subsistence farming – food production for home consumption – the optimal choice (see e.g. De Janvry et al. 1991; Jayne, 1994; Fafchamps, 1999; Key et al., 2000; De Janvry and Sadoulet, 2006)². Widespread subsistence farming leads to low productivity and low growth in agriculture. And since developing countries have large agricultural sectors with a comparative advantage vis-à-vis non-agricultural sectors, potentially large multiplier effects from agriculture to the remaining sectors of the economy and few alternative growth strategies (de Janvry and Sadoulet, 2010), a stagnant agricultural sector is likely to obstruct economic growth of these countries.

The question arises how can countries overcome this subsistence trap? A possible way out of this trap is to reduce transaction costs for smallholders. Transaction costs – costs incurred in order to sell on the market – include costs of information, collection, loading and transport of goods, bargaining on prices and conditions, monitoring and insurance, with transport costs usually considered to be the largest component. It is often claimed that transaction costs, which are only partly observed, are large and the key cause of not selling on the market (De Janvry and Sadoulet, 2006). Conversely, improved access to markets – both the mere existence of physical markets, but also institutions and a logistical and marketing infrastructure – decrease transactions costs and should, thereby, trigger smallholders to cultivate crops for the market. Transaction costs, hence, play an important role in explaining the cash crop – food crop decision.

¹ Food crops may also be sold on the market and, hence, are not necessarily or exclusively used for subsistence.

² Promotion of either food crops or commercial crops is also at the heart of policy discussions on economic growth and development (see e.g. Harrigan, 2003, 2008).

The current paper aims to contribute to this literature by investigating tobacco production in Malawi. In particular we study the impact of improved market access – caused by the introduction of a new auction floor – on the household decision to grow tobacco. Tobacco in Malawi is by far the most important cash crop and export crop, grown in nearly all Malawi districts and, by regulation, exclusively sold on auctions floors. All costs associated with selling tobacco through auctions are on account of tobacco growers. These transaction costs are substantial: estimates (2000/01) are in the range of 14.5% to 22.5% of sales value (see FAO, 2003). In 2004 an additional auction floor started operations in Chinkhoma, Kasungu district, on top of the three already operational auction floors (in Limbe, Blantyre; Kanengo, Lilongwe and Mzuzu, Mzimba). Due to a reduction in transport costs the new auction floor offered farmers in its neighbourhood an opportunity to produce tobacco in a commercially viable way. We exploit the introduction of this new auction floor to quantify the impact of a reduction in transport costs on tobacco area productivity (the intensive margin) and tobacco crop area (the extensive margin). For the empirical measurement of impact we make use of aggregate annual area and production data of smallholders at Extension Planning Area level (EPA), in total 198 EPAs, for a period of seven years, from crop seasons 2003/04 to 2009/10, covering the whole of Malawi.

The paper is organised as follows. In Section 1 we position this study in the literature and highlight its contribution. In Section 2 we describe the Malawi tobacco industry: its importance for the Malawi economy, the transition from estate based to smallholder based tobacco production in the 1990s and the marketing institutions in the tobacco commodity chain. In Section 3 we show how we measure the impact of improved market access for tobacco smallholders in Malawi. In Section 4 we present and discuss the estimation results. In Section 5 we consider alternative explanations and potential threats and run robustness checks. Section 6 summarizes the main findings.

1. What does the literature tell?

What causes farmers to grow low yielding food crops for home consumption rather than high return cash crops for the market? And what explains that large groups of farmers prefer not to participate in the market? Various researchers have modelled the decision to grow either subsistence crops or cash crops, and the decision to participate in the market (De Janvry et al.,1991; Goetz, 1992; Jayne, 1994; Omamo, 1998; Key et al., 2000; Renkow et al., 2004; de Janvry and Sadoulet, 2006). Due to transactions costs households have a tendency to get trapped into self-sufficiency and limited participation in the market explains a sluggish supply response (De Janvry et al.,1991). The wedge between producer prices for home produced maize and consumer prices for maize purchased in the market drives the decision to cultivate food crops rather than cash crops, and this wedge is especially large in rural areas, requiring a large decrease of consumption prices to make cash crop production attractive (Jayne, 1994). Transport costs between farms and markets alone are sufficient to account for observed food dominated cropping patterns as optimal responses (Omamo, 1998). The mutual dependence between food crop and cash crop cultivation is similar to non-separable household decisions with incomplete markets (see e.g. De Janvry and Sadoulet, 2006). In a generalization of the model proposed by Goetz (1992), Key et al. (2000) show that both proportional and fixed transaction costs matter: supply response to a price increase is partly due to producers who enter the market (60%), and partly due to those producers who are already sellers on the market (40%). Several researchers investigated the implications of high transaction costs empirically (e.g. Fafchamps and Vargas Hill, 2005, and Minten and Kyle, 1999; Jacoby and Minten, 2009). In choosing between selling to an itinerant trader at the farm gate or carrying output to the nearest market town, farm households are more likely to sell to the market when the quantity sold is large and the market is close by, and wealthy farmers are more likely to

travel to distant markets (Fafchamps and Vargas Hill, 2005). Differences in food prices between producer regions and urban areas are explained by transportation costs, and road quality is the key determinant of transportation costs (Minten and Kyle, 1999). Large gains in income may be realised from improved road infrastructure for remote households, although still modest relative to improved non-farm earning opportunities in town (Jacoby and Minten, 2009). Another strand of empirical work focuses on the impact of search costs – another large component of transaction costs – on behaviour and market prices. Improved market information for households is shown to significantly raise the probability of participating in the market as a seller or a buyer (Goetz, 1992). Various studies exploiting the roll-out of mobile phones have investigated impact on market prices and behaviour (see for example Aker, 2010; Jensen, 2007). In a particularly relevant study on the soy market in the central Indian state of Madhya Pradesh, Goyal (2010) investigates the impact of a direct marketing channel for farmers, in the form internet kiosks offering price information and warehouses offering quality testing and direct sales to the end-user (a private company), and thereby bypassing intermediary traders. As a result, soybean prices increased, price dispersion decreased and area under soy cultivation increased.

The literature offers persuasive evidence, both theoretical and empirical, for the key role that transaction costs play in explaining subsistence farming, on the impact of transaction costs on prices, arbitrage and economic behaviour, on supply, and on the potential welfare improvements that reductions in transaction costs can generate. In the current paper we complement the work on the choice between food and cash crops, by showing empirically the importance of transport costs in supply response in cash crop cultivation. Contrary to most work (but similar to Goyal, 2010) we investigate the impact of a change in marketing infrastructure. Next, we exclusively look at transport costs, approximated with distance to auctionfloor, rather than search costs, or any other form of transaction costs. Our analysis

exploits a reduction in transport costs, and the consequent but implicit increase in farm gate prices, and investigates its impact on supply responses of farmers. We find a statistically significant impact of transport costs on crop area and production of cash crop cultivation. A reduction of transport costs is, thereby, shown to trigger commercial agriculture, and confirms results of other empirical work (Key et al. 2000; Renkow et al., 2004; Jensen, 2007). The empirical investigations also fit the literature that seeks to reveal constraints to export growth strategies and to highlight the potential of export led growth strategies in poverty alleviation (see e.g. Balat et al., 2009).

2. The Malawi tobacco industry

Various articles and publications describe market developments in the Malawian tobacco sector, the evolution of the marketing and regulatory infrastructure over time and the transformation that took place from the colonial estate based production to smallholder based production since the end of the 1980s (see Kydd and Christiansen, 1982; Orr, 2000; Diagne and Zeller, 2001; Jaffee, 2003; World Bank, 2004; Poulton et al., 2007; Tchale and Keyser, 2010). We draw on these sources to highlight the key developments and institutional changes relevant to the empirical analysis. Complementary to those descriptions we analyse aggregate historical auction data (1960-2010, source: TCC), the complete 2009 auction transaction data for all tobacco auction floors (source: TCC) and tobacco growers information in the Malawi LSMS household data (IHS 2 (2004/05), IHS3 (2009/10) and IHPS (2013); source: NSO).

The role of tobacco in the domestic economy of Malawi

Tobacco is, by far, the most important export product of Malawi accounting for a share of 45% to 65% of total merchandise exports (1994 to 2009, NSO data). The second largest single export product (tea or sugar) account for only a small fraction of total merchandise exports. Tobacco exports also account for about 60% of foreign exchange earnings and, as

result of this, the Malawi kwacha - US\$ exchange rate tends to fluctuate with export revenues from tobacco (and the tobacco season). All tobacco is exported: Malawi does not have a domestic cigarette industry. The direct contribution of tobacco to GDP, measured as the export value of tobacco in terms of GDP, varies from 9% to 16% (1994 to 2009, NSO data).

Tobacco is cultivated by 19% of the smallholder households, around 375,000 (2004). The bulk of the tobacco growing households – around 65% – are poor or very poor (Economic Council (2000)). In the period from 2003 to 2010 aggregate smallholder crop area allocated to tobacco varied from 141,000 to 184,000 hectares, and smallholder crop production from 95 to 208 thousand tons (source: Agro Economic Survey, Ministry of Agriculture and Food Security). Using a methodology employed by the FAO (FAO, 2003), direct employment in tobacco production and marketing (including processing, transport, auctioning and research) varied from 11% to 19% of total labour supply during 2000-2009.

Tobacco exports generate a major contribution to total government tax revenue in the form of withholding tax levied at the auctions, together with export taxes and export surrender requirements imposed by the Reserve Bank of Malawi. All tobacco taxes and levies add up to an estimated share of total government tax revenue of 30% in 2000 and around 20% in 2008. Jaffee (2003) reports 23% and FAO (2003) writes: "...tax accounted for more than 20 percent of total national tax revenue". The large share of tobacco proceeds that flows to the Government of Malawi makes the government a major stakeholder in the tobacco industry. Related to this a variety of rents may arise due to lack of competition, lack of transparency and lack of accountability (see e.g. Koester et al., 2004).

In summary the figures indicate that tobacco is of extraordinary importance to the Malawi economy. The role of tobacco may extend well beyond these figures, due to indirect effects, backward and forward linkages and dynamics. Some authors claim Malawi's export of tobacco to be the major driver of economic growth (see e.g. Lea and Hammer, 2009).

Tobacco cultivation in Malawi: from colonial heritage to smallholder domination

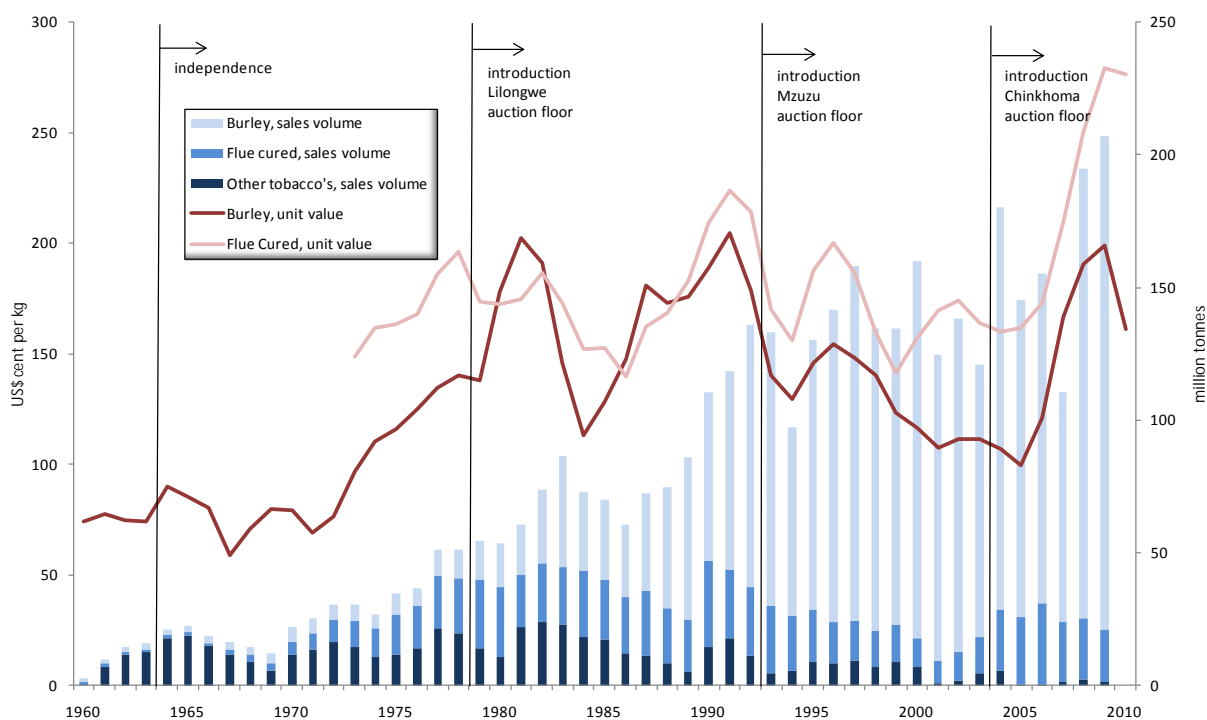
The Special Crop Act of 1964 had created a dual sector, with special privileges for estates and with restrictions for smallholders. Urged by donors to implement liberalisations in the tobacco industry, the 1993 newly elected government introduced amendments to this 1964 Special Crop Act that allowed smallholders to grow burley tobacco (Jaffee, 2003)³. The Act was fully repealed in 1996, which included the abolishment of special marketing rights to estates. By 1996/97 all restrictions for smallholders to grow and market tobacco were removed (Diagne and Zeller, 2001). In the course of the 1990s, the change in regulation have given rise to a complete transformation from estate based tobacco cultivation with a high share of western type tobacco's, to a smallholder based tobacco cultivation with a high share of burley tobacco (see Figure 1). High profitability of tobacco as a cash crop – the only really remunerative cash crop available to smallholders – and the wide-spread technical knowledge on tobacco cultivation – since many farmers worked previously on estates as labourers – triggered high growth of smallholder tobacco production. The increase of smallholder production was accompanied by the formation of burley clubs and the introduction of intermediate buyers who provided the logistical link from farmers to auction floors and access to these auction floors (FAO, 2003). Credit for tobacco growers was made available by the Malawi Rural Finance Company (Jaffee, 2003). The transition to a smallholder based tobacco cultivation with a high share of burley tobacco is clearly visible in the data. Over the years, aggregate sales volume at auction floors – a reasonable indicator of aggregate Malawi tobacco production⁴ – shows a nearly continuous upward development (see Figure 1), almost entirely

³ Burley tobacco is a light air-cured tobacco used primarily for cigarette production. Western type tobacco's grown in Malawi are Flue-cured tobacco (also Virginia), NDDF and SDDF (respectively Northern and Southern Division Dark Fired). These latter types are smoke and fire dried and aged in curing barns, and thereby more capital and processing intensive relative to burley tobacco.

⁴ Since all tobacco exported from Malawi is required to be sold at auction, unit values and sales volume at auctions are reasonable indicators of average Malawian market prices and aggregate production, despite small quantities of tobacco sourced from Zambia and Mozambique or illegally exported (see e.g. Koester et al., 2004). Note that

due to burley tobacco, with high growth rates of burley tobacco since the end of the 1980s, which slightly levelled off by the end of the 1990s.

Figure 1 Auction Sales Volume and Unit Values of Burley and Other Tobacco^a



^a Note to figure – Nominal unit values in US\$ cent per kg are on the left axis and sales volume in million tonnes on the right axis. Other tobacco's produced in Malawi are NDDF, SDDF (resp. Northern and Southern Division Dark Fired, so-called western tobacco's) and Sun Air; source: annual aggregate data from the Tobacco Control Commission, Malawi.

Figure 1 further illustrates the parallel development of burley and flue cured auction unit values, with the latter in most years higher. Visual inspection of the figure suggest that (lagged) prices move jointly with production in a more or less systematic way, reflecting a positive response of production to auction prices (see also Jaffee, 2003 and our own estimates).

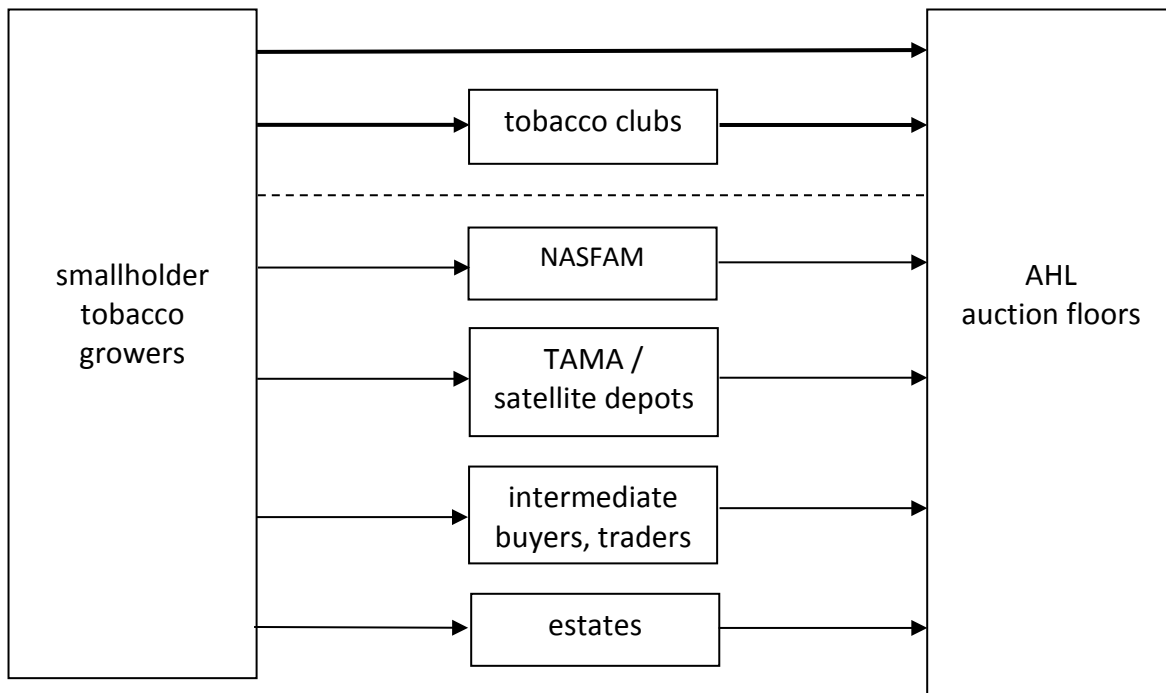
Tobacco marketing: auctions, regulations, farmers clubs and other institutions

The tobacco auction system in Malawi has a long history, which dates back to the colonial times, at the start of the 19th century and which was, for a long time, adapted to production and

empirical estimations in this study are based on data from the Ministry of Agriculture and Food Security (Agro-Economic Survey) and not from the Tobacco Control Commission.

marketing needs of estates. We focus on the period since the 1990s, when the tobacco sector was liberalized. Transport of tobacco to auctions was – both pre and post liberalization – on account of tobacco farmers. Hence, starting in the 1990s a logistical infrastructure for tobacco transport and marketing from rural areas to auctions was emerging to service smallholder farmers. Of key importance in this context are farmer clubs or burley clubs: groups of 10 to 30 farmers that share specific services. Upon registration with the Tobacco Control Commission (TCC) clubs are allocated a quota and are entitled to receive burley seed, fertilizer, advice on cultivation and extension support. From 1991/92 onwards clubs are authorized to sell directly on the auction floors and, since 1994, also to intermediate buyers, introduced in 1994 also to help smallholders to transport their burley tobacco to the auction floors (Orr, 2000). By 1996/97 smallholders were, for the first time, allowed to produce and market tobacco without any restrictions (Diagne and Zeller, 2001). In general, the process of liberalisation has spurred market access for smallholders: by 1996 83% of smallholder tobacco was marketed directly to the auction floors (Diagne and Zeller, 2001). Access to auctions and thereby access to world market prices, credit facilities and economies of scale in transport are the major incentives for smallholders to join a burley club (Orr, 2000; Negri and Porto, 2008). From 2000 to 2010 the number of registered burley clubs nearly tripled from around 20,000 at the start of the 2000s to close to 60,000 in 2010, of which more than half is registered at the Lilongwe auction floor (source: TCC). The existing Tobacco Association of Malawi (TAMA) and the National Association of Smallholders Farmers of Malawi (NASFAM), which was established in the 1990s, also assisted in the organisation, collection, storage, transport and sale of smallholder tobacco from rural areas to the auction floors. Shortcomings to the marketing infrastructure – which is continuously developing – were experienced in the area of widely divergent transport rates, storage losses and lack of accountability (see Jaffee, 2003).

Figure 2 Trade channels for smallholder tobacco growers



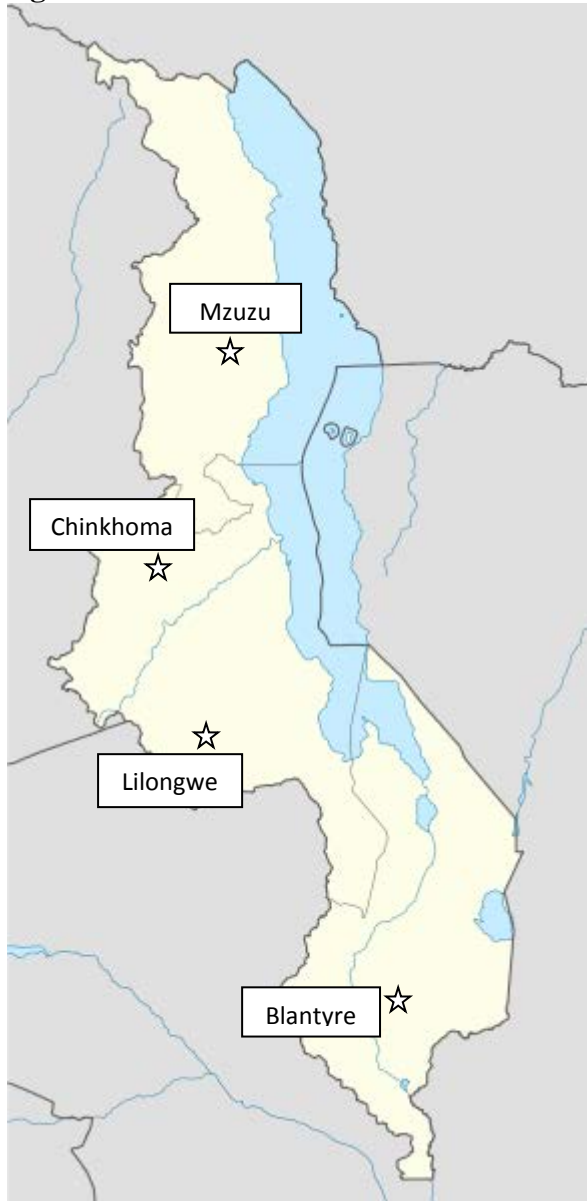
Source: author

Important for our empirical work is that transaction costs of tobacco (including transport costs) are on account of smallholders. Evidence supports that most tobacco growing households sell their tobacco directly on auction floors (Diagne and Zeller, 2001: 83% and IHS2 (2004/05): 85%). Likewise transaction costs are on account of smallholders. Consequently, net tobacco revenues received by tobacco growers increase if transaction costs, the costs associated with selling tobacco on an auction floor, are reduced.

Tobacco marketing is regulated by the Tobacco Control Commission (TCC), a statutory governmental body. The TCC is responsible for market regulation and control, licensing of farmers, quality standards, data & statistics of the tobacco sector, and advising the government on tobacco issues. Operations on all tobacco auction floors are run by a single private sector company, the Auction Holdings Limited (AHL). The establishment of an auction floor requires complementary investments from buyers to properly organize after

sales processing, storage and international transport: this makes the auction floor location decision partly dependent on the investments of buyers.

Figure 3 Tobacco auction floors in Malawi



Source: author

As early as 1939 tobacco was auctioned only at the Limbe auction floor, near Blantyre in the south of Malawi. In more recent years the centre of tobacco production moved in northern direction: auction floors were established in 1979 in Kanengo, near Lilongwe in Central

Malawi; in 1993 in Mzuzu in Northern Malawi; and in 2004 in Chinkhoma in the central district Kasungu, between Lilongwe and Mzuzu (see Figure 3). Auction floors normally open from mid-March and close towards the end of October. According to weekly reports from TCC⁵ direct trade and contract trade is primarily important for specialty tobacco's (Flue Cured, NDDF and SDDF) and plays a negligible role for the marketing of burley tobacco.

A limited number of companies is active on the demand side. Most companies are subsidiaries of large international traders or international cigarette manufacturers. Over the years the composition of the buying side has changed due to new entrants, mergers and takeovers. However, concentration on the buyer side remains high. The presence of only a limited number of buyers on the auction floors (7 in 2011) raises suspicion of a lack of competitiveness of tobacco pricing and collusion between buyers at the auction floors (see e.g. Koester et al. 2004, Otañez at al., 2007). This is particularly manifest with occasional outbursts of protest from tobacco farmers who complain about the low prices at the auction.

3. Behaviour, data and data sources, and empirical strategy

How do transaction costs influence farmers' behaviour?

The theoretical background of the impact of (a reduction in) transaction costs on crop area and production derives from the model of rural household behaviour under market failures (Goetz, 1992; Key et al, 2000; De Janvry and Sadoulet, 2006) which we closely follow. Households maximise the expected present value of current and future utility, where utility is determined by consumption and preferences. In formula: $I \doteq \sum_{t=0}^{\infty} \beta^t U^c(z_t, V_t)$ where β is the rate of discount, u is utility, c is consumption, z are preferences and subscripts t is time. For each year the households choice variables are consumption, production, inputs, marketed surpluses and savings. Maximization is subject to a budget constraint where the budget

⁵ These reports are only available for the period from 2001 to 2006.

constraint, again for each period, is the balance of value of marketed surpluses (both positive and negative) over all goods, fixed transaction costs, transfers and savings. Marketed surpluses, both positive and negative ones, are evaluated at their relevant prices, where purchase and sales prices are assumed to be different, and include variable transaction costs. In formula the budget constraint for a tobacco seller and a (net) maize buyer household is:

$$P_{t,c}^a L_{t,c}^a + P_{t,c}^b L_{t,c}^b + \dots + P_{t,c}^n L_{t,c}^n + P_{t,c}^m L_{t,c}^m + P_{t,c}^f L_{t,c}^f + P_{t,c}^g L_{t,c}^g + P_{t,c}^h L_{t,c}^h + P_{t,c}^i L_{t,c}^i + P_{t,c}^j L_{t,c}^j + P_{t,c}^k L_{t,c}^k + P_{t,c}^l L_{t,c}^l + P_{t,c}^m L_{t,c}^m + P_{t,c}^n L_{t,c}^n + P_{t,c}^o L_{t,c}^o + P_{t,c}^p L_{t,c}^p + P_{t,c}^q L_{t,c}^q + P_{t,c}^r L_{t,c}^r + P_{t,c}^s L_{t,c}^s + P_{t,c}^t L_{t,c}^t + P_{t,c}^u L_{t,c}^u + P_{t,c}^v L_{t,c}^v + P_{t,c}^w L_{t,c}^w + P_{t,c}^x L_{t,c}^x + P_{t,c}^y L_{t,c}^y + P_{t,c}^z L_{t,c}^z + C + G + Q = 0$$

where L^a is market price, P^b is proportional transaction costs, P^i is fixed transaction costs, I^i is market surplus (negative in case of maize) and subscripts are crop (tobacco and maize) and time. (The dots (...) stand for the n other crops and goods, which are ignored for the sake of simplicity). A key and extensively studied property of this household model, so-called non separability, is that household responses in one market are affected by failures in other markets. An increase in cash crop prices can only lead to reallocation of land and labour of a household from food crops to cash crops if the increased revenues offsets the decrease in food production and the increased costs of the household on purchased food. This explains low supply response in case of market failures and, hence, also in case of high transaction costs. Alternatively, a productivity increase in food production or a drop in prices of imported food may free up land, labour and other inputs for cash crop production.

With both variable and fixed, household specific transaction costs it is clear that sales and purchase prices for each good are also household specific. If households rely for food on own production and produce a cash crop for exchange on the market – reasonably alike tobacco smallholders in Malawi who also cultivate maize – the formal solution to the model implies a shadow price of the home consumed crop (I_a^U), or in formula:

$$(M_a^U, L_b^U, L_a^U, V_a^U) = ?(L_a^U, L_o^U, U^U, V_o^U) \text{ where } U^U \text{ is household income which is written as: } U^U = L_a^U M_a^U + L_b^U M_b^U + L_a^U M_a^U + G.$$

Participation in cash crop market depends on the

Distance from EPAs to the different auction floors is road distance, since road distance is the relevant concept for transport costs. All distance are sourced from Google Maps⁷. Potential measurement error may arise because latitude-longitude coordinates for identification of EPAs – usually the main town / village in the EPA – will not necessarily coincide with the tobacco area in the EPA. This measurement error is correlated with the size of the EPA.

Apart from using EPA data of a few additional crops (maize, groundnuts and pulses), there is, unfortunately, no further systematic information at EPA level available. For the analysis we have complemented and matched the EPA data with data from various other sources, in particular rainfall, prices, population and various distances. These data have different levels of aggregation and varying missing observations, which introduces a certain degree of measurement error. Crop prices are, for example, available for respectively 50 to 70 markets and in varying degrees of completeness. Rainfall data are available for around 30 weather stations, but, luckily, without missing observations. In summary, we have data on crop prices, production and area of tobacco and alternative crops, total crop area, rainfall, population, land area and various distances to analyse dynamics in tobacco cultivation.

Statistics of these variables are shown in Table 1: most variables differ between intervention and control group, which is, apart from the price data, merely a reflection of the differences between EPAs. Maize prices are similar, which is an interesting exception. Rainfall is slightly higher in the intervention EPAs. Spatial and population variables indicate that the intervention EPAs are somewhat more remote from cities and towns, less integrated in a network and with lower population densities. Further details on data, data sources and variable construction are in Appendix A.

⁷ We have assumed no change in road distance due to maintenance or road construction.

Table 1 Summary of data for estimation

Variable	intervention		control		t
	mean	standard deviation	mean	standard deviation	
rural / urban					
consumer price index	-	-	-	-	-
market level data					
tobacco price (MK)	71.6	34.0	67.5	28.3	2.2
maize price (MK)	22.3	7.9	22.2	8.1	0.2
groundnuts price (MK)	146.2	51.1	156.0	46.3	3.2
weather station level data					
rainfall (mm)	981.6	298.4	929.2	242.3	3.0
EPA level data					
crop cultivation					
tobacco area (1000ha)	2.0	1.4	0.5	0.7	25
maize area (1000ha)	10.8	5.4	7.7	3.8	11
groundnuts area (1000ha)	2.3	1.6	1.1	0.9	16
pulses area (1000ha)	3.1	2.3	3.0	3.1	0.1
total crop area (1000ha)	20.6	8.8	15.5	7.6	9.7
tobacco production (ton)	1698	1407	474	722	20
maize production (ton)	19769	13102	11660	7642	13
groundnuts production (ton)	2195	2098	926	1001	14
pulses production (ton)	1930	1774	1866	2251	0.4
Population					
population density (people/km2)	138.1	59.9	204.2	179.3	6.4
Geography					
distance to town (km)	39.2	22.4	36.2	19.7	2.2
distance to city (km)	88.6	28.6	66.1	46.5	7.8
agglomeration index	13.7	4.1	17.9	10.6	6.9

Note to table: Observations of data at the rural-urban level, at the market level, and at the weather station level pertain to the number of independent observations in the original data set. Mean, standard deviation and t test for all variables are, however, calculated on the basis of EPA level data or the EPA level variants of market level data and weather station level data.

Intervention locations

Tobacco farmers that benefit from the introduction of the new auction floor in Chinkhoma in 2004 are identified by determining the minimum of the distances from each EPA to the different auction floors. If in 2004 the Chinkhoma auction floor has become the closest auction floor, tobacco growers in those EPAs have realised a reduction transport costs to the auction floor. Practically this implies that all EPAs in the districts Kasungu and Nkhotakota, a large part of locations in the districts Ntchisi, Dowa and Mchinji, and a few in the district Mzimba are intervention locations. In all this concerns 31 EPAs / locations, 15.3% of all locations.

The distribution of sales by district of origin shows that the Chinkhoma auction floor also attracts tobacco outside these districts (e.g. from Lilongwe, Rumphi and Salima district;

Estimating the impact of a change in distance to the tobacco auction floor

The objective of this investigation is to measure supply response due to a change in market access. More specifically we measure if and to what extent the new auction floor has given rise to changes in area (the extensive margin), production and yield (the intensive margin). To this end the following equation, equivalent to a difference-in-difference equation, is used to estimate the impact of the introduction of a new auction floor in Malawi (the intervention):

$$\ln y_{it} = \beta_0 + \beta_1 \ln(\text{distance to auction floor}_{it}) + \beta_2 \text{area}_{it} + \beta_3 \text{production}_{it} + \beta_4 \text{yield}_{it} + \beta_5 \text{covariates}_{it} + \beta_6 \text{region}_{jt} + \beta_7 \text{EPA}_{it} + \beta_8 \text{year}_{t} + \beta_9 \text{EPA} \times \text{year}_{t} + \epsilon_{it}$$

where y_{it} is the outcome indicator and stands resp. for area, production and yield of EPA i in year t , covariates_{it} is a set of k exogenous or predetermined covariates, region_{jt} is a dummies for region j , β_0 and β_1 are EPA fixed and year effects, and ϵ_{it} is an error term with zero mean and constant variance. Outcome and intervention variables, and exogenous and predetermined covariates are all transformed into natural logarithms to allow for non-linearities in variables. The coefficient of the variable *distance to auction floor* (β_1) the parameter of interest, measures the average impact of changes in *distance to auction floor* on the dependent variable. We use the above specification as the basic specification of our estimations.

Covariates (covariates_{it}) derive from profit maximizing behaviour of farmers. Expected profits are determined by expected yield, and expected input and output prices. Hence, both tobacco area and production are likely to respond to expected prices of tobacco (see e.g. Figure 1 in this study, but also Jaffee, 2003) and expected prices of alternative crops. Therefore, we use previous year farm gate tobacco prices and previous year farm gate prices of other crops (maize, the major alternative crop, but also groundnuts and/or pulses), since tobacco area and production decisions are guided by tobacco prices in the first place, but are also made jointly with decisions on alternative cultivation options. We use farm gate prices since these prices –

in contrast with tobacco auction prices – exclude transaction costs and drive smallholder decisions⁸. Given missing markets and non-separability of household behaviour cash crop decisions are also likely to respond to food crop developments like yield (De Janvry and Sadoulet, 2006). This justifies to include lagged yield of food crops as covariates. Additionally, like in most rain-fed agriculture, output of tobacco cultivation is determined by rainfall. Therefore we employ current rainfall as a covariate in tobacco production (but not in tobacco crop area). Finally, we experiment with several geographical variables to approximate other transaction costs, access to markets and other agglomeration effects, like population density (for within-EPA transaction costs and agglomeration effects), distance to city or town, and a distance weighted city and town size index (for outside-EPA transaction costs and agglomeration effects). Data sources and data construction are explained in detail in Appendix A.

Using generalised propensity scores and dose response functions for identification of impacts

The location of the new tobacco auction floor is not randomly assigned. The auction company will have carefully considered alternatives and investigated the optimal location for doing this investment, basing its eventual choice on an assessment of current and expected turnover of tobacco and long run profit potential of auction services at different (hypothetical) locations. Consequently, causality may not run (only) from market access to decisions of tobacco growers to grow tobacco, but also the other way around, from (expected) tobacco production to the establishment of an auction floor. As a result, OLS estimations are potentially biased: estimates may not reflect the isolated impact of a change in transaction costs.

We propose a propensity score method to reduce the potential bias in the estimations: unfortunately, propensity score matching models are mainly developed for binary treatment

⁸ Because of its constructed nature (see appendix A) we are reluctant to use the farm gate tobacco prices as direct evidence for the impact of the reduction in transport costs caused by the new auction floor. The results of an event plot are, however, quite supportive (see Appendix C).

variables and hence not suitable for our continuous *distance to auction floor* treatment variable. An extension of the propensity score method for a continuous treatment setting is developed in Hirano and Imbens, 2004, and will, hence, be used for the empirical work in this study. We briefly summarize this method and the associated STATA commands (see Bia and Mattei, 2008), with particular attention to the underlying assumptions of the treatment variable and testing the balancing property. The continuous treatment corollary to binary treatment propensity scores is the Generalised Propensity Score (GPS). Likewise, the GPS can be used to eliminate biases associated with differences in the covariates. The propensity function is defined as the density of actual treatment conditional on covariates ($r(t,x) = f_{T|X}(t,x)$) and the Generalised Propensity Score is $R = r(T,X)$. The GPS balancing property requires that within strata with the same value of the function $r(t,x)$, the probability of a specific value of the treatment (t) should be independent of the covariates (X). Hence, if we estimate the GPS $r(t,x)$ within a stratum, it should be the same for different values of t .

The function that maps the relationship between outcome and continuous treatment is labelled the (unit-level) dose response function ($Y_i(t)$): we are interested in the average dose response function. The dose response function is estimated at a particular level of the treatment as the average of the conditional expectation over the GPS at that particular level of the treatment, $E[Y_i(t) | R = r(t,x)]$. Although Propensity Scoring techniques are only a modest improvement over OLS, GPS and the dose-response estimation that employs GPS, does offer more flexibility in functional form (see also the empirical section).

4. Estimating impact of market access on tobacco area, production and yield

Selected estimation results for our basic specification are reported in Table 2. The table reports estimations results for a basic specification, containing EPA and year fixed effects (column 1, 3 and 5), and a specification additionally including covariates (column 2, 4 and 6).

Unlike the event estimations from the previous section, that clearly indicated varying and lagged responses, the specification imposes a fixed impact over the years without lag⁹. Additionally a relatively broad definition of intervention EPAs is chosen, including EPAs of which the distance to the new Chinkhoma auction floor comes close to but is not the minimum distance to auction floor. A broader definition corresponds with the source of transactions in the 2009 transaction data (see Appendix F). With respect to covariates we experimented with trends, rainfall, and lagged tobacco and maize prices. Some of these variables (rainfall, tobacco price) will only affect production and much less crop area.

Coefficients of the *distance to auction floor* variable in the production equation (columns 3-4) are negative and strongly significant. This outcome supports the hypothesis that a larger distance to the auction floor increases transaction costs and thereby decreases commercial attractiveness to cultivate tobacco. Due to the logarithmic transformation, coefficients may be interpreted as elasticities: The elasticity of production with respect to distance to auction floor is close to 1: a 10% increase in the distance to the auction floor decreases production, also with 10%. In the crop area estimations (columns 5-6) the coefficients of the *distance to auction floor* variable is also negative and significant (at the 10% level). Hence, and likewise, a shorter *distance to auction floor* decreases transaction costs and thereby increases commercial attractiveness of tobacco cultivation, leading to a larger allocation of crop area to tobacco cultivation. The size of the elasticity, however, is much lower: a 10% decrease in the distance to auction floor increases tobacco crop area only with 4%. Estimations on yield (columns 1-2) show negative and highly significant coefficients of *distance to auction floor*, consistent with the production and area results¹⁰.

⁹ We experimented with various lag structures but this exercise did not offer a clear alternative.

¹⁰ Similar to the event estimations, we have estimated impact using a binary impact variable, a variable that takes the value of 1 if the distance to auction floor decreases and zero elsewhere. The estimations measure the average impact of the reduction in distance to auction floor and indicate increases in production, crop area and yield of respectively 33%, 16% and 22%, all statistically significant (results available from the author).

Table 2 Market access in tobacco: basic specification and covariates

dependent variable	ln(production)		ln(area)		ln(production per ha)	
	(1)	(2)	(3)	(4)	(5)	(6)
ln(distance to auction floor)	-0.937*** (0.341)	-1.020*** (0.342)	-0.376* (0.226)	-0.404* (0.225)	-0.524*** (0.140)	-0.597** (0.293)
trend		0.077*** (0.028)		0.078*** (0.013)		0.023*** (0.007)
ln(rainfall)		0.202* (0.116)				
ln(tobacco price), lagged		0.085 (0.091)				
number of observations	1152	1148	1152	1148	1152	1148
R ²	0.939	0.944	0.958	0.959	0.517	0.563

Notes – The table reports ATE of distance to auction floor on tobacco area, production and yield. Estimations are based on annual data from 2003-04 to 2009-10 (seven years). All estimations include EPA and year fixed effects. All equations are estimated with OLS. Robust standard errors clustered by EPAs are given in parentheses (.) below the coefficient. R² = coefficient of determination. * means significant at the 10% level ($p < 0.10$), ** at the 5% level ($p < 0.05$), *** at the 1% level ($p < 0.01$).

Coefficients of covariates have expected signs and are on some cases statistically significant: Rainfall is statistically significant in the production estimations, while lagged tobacco price is has the expected sign in the production equations, ut is not really significant. An included trend variable is statistically significant in all three equations. The key message concerns the impact variable: coefficients of *distance to auction floor* are statistically significant and with the same sign as in the basic specification and estimations are, hence, reasonably robust for inclusion of covariates.

On the basis of these estimations we infer – and this is a key outcome of this work – that a reduction in transport cost leads to an increase in crop area and production. Under the assumption that *distance to auction floor* is a good approximation of transport costs, a 10% reduction in transport cost will lead to an increase in crop area of around 4% and an increase in production of around 10%. The impact of a reduction in transport costs is larger on production than on crop area and, as a result, a decrease in distance to the auction floor will also increase yield (production per hectare). Apparently there is sufficient scope to increase production and to increase production per hectare, before productivity hits its limits, because

more and improved inputs, further improvement of cultivation practices, and more intensification of cultivation does not yield higher production. If that point is reached production per hectare may decrease as a result of cultivation of less productive crop land or as a result of less experienced farmers engaging in tobacco cultivation. We may conclude that the impact of a reduction of the distance to auction floor – a reduction in transport costs – runs through the intensive margin (an increase in production per hectare: an increase in production jointly with a smaller crop area increase), and not through the extensive margin (a decrease in production per hectare, an increase in production jointly with a larger crop area increase).

A supply response along the intensive margin is associated with making steps on the learning curve, increasing inputs like fertilizer and irrigation, stepping up crop maintenance, or, more in general, intensification of tobacco cultivation. Such a development is probable to take place on well-to-do farm households that improve on their already existing tobacco cultivation. Increases in area productivity are costly and therefore often beyond the opportunities of poor households. Conversely, a supply response along the extensive margin would be associated with extension of crop area, both by existing tobacco growers and by new growers, to area with less suitable soils and poorer cultivation conditions. New growers also need to learn about cultivation practices. Both will lead to lower production per hectare. Hence, a supply response along the extensive margin is more compatible with responses of poor smallholders and is thus likely to lead to poverty alleviation and improved welfare. Unfortunately the data at hand do not support such a conclusion. We may, nevertheless, consider the results good for poverty alleviation since all data used for the estimations concern smallholders.

Using generalised propensity scores in the estimation of dose response functions

In order to reduce the potential bias that arises with OLS because of selectivity of the intervention locations, we employed a generalised propensity score method (GPS), which is specifically developed to generate unbiased estimates of the population Average Treatment Effect (ATE) for the case of continuous treatment (see Hirano and Imbens, 2004)¹¹.

Table 3 Market access in tobacco: dose response estimations

Dependent variable:	ln(production)	ln(area)	ln(production per ha)
	(1)	(2)	(3)
ln(distance to auction floor)	-0.007*** (0.0014)	-0.007*** (0.0014)	0.0001 (0.0003)
generalized propensity score (GPS)	-0.990*** (0.359)	-0.744** (0.349)	-0.246*** (0.070)
number of observations	1038	1038	1038
F (.)	(2,1035) 14.43	(2,1035) 14.58	(2,1035) 6.80
Prob > F	0.000	0.000	0.001
Adjusted R ²	0.0253	0.0255	0.0111

Notes – The table reports dose response estimations of the introduction of a tobacco auction floor on tobacco production per hectare, using generalized propensity scores, the continuous treatment variant of propensity scores. Covariates used in constructing the generalized propensity score are: rainfall, real lagged tobacco and maize price, lagged area productivity in maize and tobacco and the lagged share of maize in total crop area. The selection of variables underlying the estimation of the propensity score support the balancing property. Estimations are based on annual data from 2003-04 to 2009-10 (seven years). * means significant at the 10% level ($p < 0.10$), ** at the 5% level ($p < 0.05$), *** at the 1% level ($p < 0.01$).

In the estimation of the GPS we have divided the intervention variable – *distance to auction floor* – into three intervals, equally sized in terms of observations. The balance of the covariates is investigated by testing whether the mean in one of the treatment groups differs from one of the others or both. For this purpose each interval is blocked into four blocks and GPS scores are compared. Covariates used in constructing the generalised propensity score are: rainfall, real lagged tobacco and maize price, lagged area productivity in maize and tobacco and the lagged share of maize in total crop area. The selection of variables underlying the estimation of the propensity score (and the selection of intervals and blocks

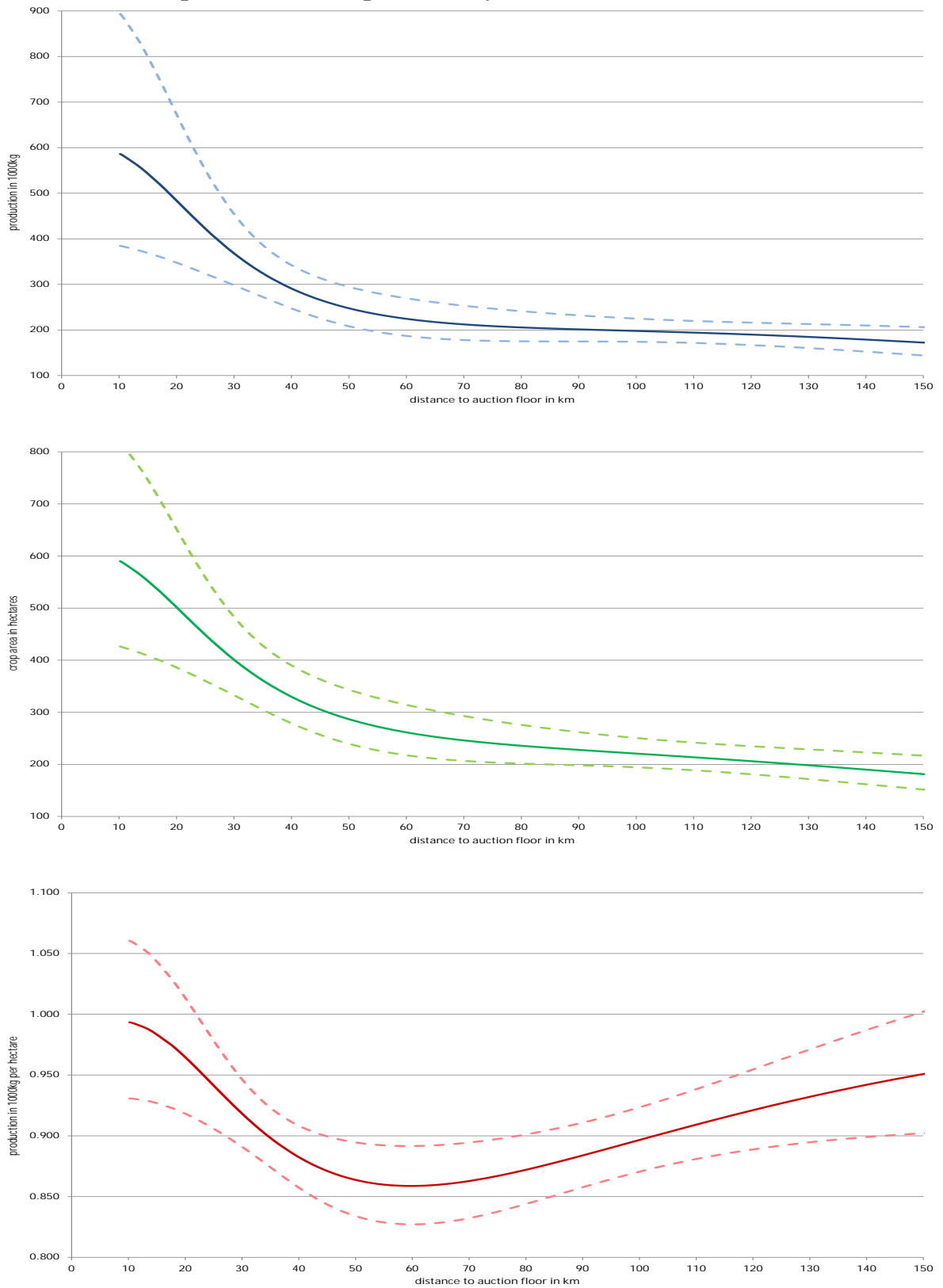
¹¹ We employed the STATA commands *gpscore* and *doseresponse*. Output reported in this section is based on these commands.

within intervals) support the balancing property at the 20% level. Bootstrap methods are used to derive standard errors and confidence intervals, using 50 bootstrap replications.

The estimations, reported in Table 3, document the results of the estimation of the dose-response function based on the generalised propensity score. The interpretation of the coefficients in the table is complicated due to reduced form nature of propensity scores and the use of a flexible specification of the dose-response function. However, the figures of the dose response function (Figure 4) are easier to digest. These figures suggest that both crop area and production are negatively correlated with distance to auction floor: a higher distance to the auction floor is associated with lower crop area and production. This confirms our difference-in-difference estimates. Moreover, the figures suggest that crop area and production in EPAs with a short distance to auction floor (less than 60 km) have a much larger response than EPAs are farther away from auction floors. In fact, the figures indicate that a change in distance to auction floor beyond 60 km triggers very little response.

Within a distance of 60 km from an auction floor, a 10 km reduction induce a 80 hectare increase (a 90 ton increase in tobacco production), while a similar reduction beyond 60 km only leads to a crop area increase of approximately 10 hectare (an increase in tobacco production of around 10 ton). The non-linear impact is due to the GPS method. For *distance to auction floor* below 60km, the *distance to auction floor* elasticity for production, estimated with OLS, correspond roughly with the elasticity of the dose response estimations. The *distance to auction floor* elasticity for crop area, estimated with OLS, is somewhat lower compared to the elasticity of the dose response estimations.

Figure 5 Dose-response function: impact of distance to auction floor on production, crop area, and area productivity



5. Alternative explanations and potential threats

Impact on other crops

The statistically significant impact on tobacco area, production and yield in the EPAs that are benefitting from the newly established auction may be a coincidental outcome that applies to all crops in these EPAs. For this reason we have repeated the impact estimations using area, production and yield of alternative crops, notably maize, groundnuts and pulses. Maize is the key food crop and produced by virtually all households. Maize accounts for more than 50% of total crop area and around 60% of the Malawi food consumption diet (MoAFS and FAO). Groundnut is (partly) a cash crop like tobacco but also a food crop. Groundnut cultivation has a country wide distribution similar to tobacco and groundnuts are also an important source of income for farm households, although less important than tobacco. Cultivation of pulses is less widely distributed. With the exception of maize, these alternative crops are, like tobacco, high value crops with per kg price of 4.5 to 8 times the price of maize. Since endogeneity of the intervention locations cannot be an issue for these alternative crops, we do not need to apply estimation methods that adjust for the associated bias and apply the previously used difference-in-difference specification. Estimations results are reported in Table 4.

The table shows that the coefficients of *distance to auction floor* are small and statistically insignificant, or, in one case, only weakly significant. Hence, the estimation results do not support a systematic and statistically significant impact on area, production and yield, of maize, groundnuts and pulses. The hypothesis that the estimated impact on tobacco area, tobacco production and tobacco yield applies to other crops as well, is not confirmed by the data. This result further strengthens our claim that improved market access – by the introduction of a new auction floor – has a positive impact on production per hectare, production and area of smallholder tobacco farmers.

Table 4 Impact estimations with placebo crops

Dependent variable:	ln(production)	ln(crop area)	ln(production per hectare)
Crop: maize	(1)	(2)	(3)
ln(distance to auction floor)	-0.061 (0.093)	-0.102* (0.055)	0.041 (0.066)
number of observations	1318	1318	1318
R ²	0.866	0.936	0.787
Crop: groundnuts	(1)	(2)	(3)
ln(distance to auction floor)	0.140 (0.087)	0.066 (0.085)	0.073 (0.044)
number of observations	1318	1318	1318
R ²	0.903	0.925	0.604
Crop: pulses	(1)	(2)	(3)
ln(distance to auction floor)	0.119 (0.0727)	0.150 (0.108)	0.031 (0.066)
number of observations	1318	1318	1318
R ²	0.952	0.924	0.607

Notes – The table reports ATE of distance to auction floor on maize, groundnut and pulses production, crop areas and production per hectare. Estimations are based on annual data from 2003-04 to 2009-10 (seven years). All estimations include EPA and year fixed effects. All equations are estimated with OLS. Robust standard errors clustered by EPAs are given in parentheses (.) below the coefficient. R² = coefficient of determination and . means significant at the 10% level ($p < 0.10$), * at the 5% level ($p < 0.05$), ** at the 1% level ($p < 0.01$).

A related issue in this context is the interaction between crops. Has the positive impact on tobacco crop area (and production) given rise to a negative impact on crop area (and production) of alternative crops? Was crop area of other crops (maize, groundnuts, pulses, etc.) substituted into tobacco area? For this purpose, we employ the event study specification, shown at the start of this paper: with this specification we show if crop area (crop production) in the EPAs that benefitted from the new auction floor has changed. The results of this exercise, again, implemented for maize, groundnuts and pulses (not shown here, available from the author), suggest no statistically significant reduction of crop area of these crops and, hence, no reduction of crop area of alternative crops, in response to the introduction of the new auction floor. We further corroborated this observation by estimating how increases in total crop area and decrease in crop area of alternative crops are associated with changes in tobacco area (see Appendix D). We find that increases in tobacco area are mostly associated with increases in

total crop area, and not with decreases in area of alternative crops. Hence, the evidence does not support substitution of area of other crops into tobacco area.

Parallel trend assumption

For the difference-in-difference estimations to generate credible results it is required that production, crop area and yield of both EPAs that benefit from the new auction floor and EPAs that do not benefit, are on a parallel trend before the start of the new auction floor in Chinkhoma. It is most popular to show graphically that this parallel trend assumption is satisfied. In fact, the standard exercise to implement is similar to estimating and drawing the event plots, shown in a previous section. If the pre “new auction floor” trends of both intervention and non-intervention EPAs are the same, then the pre introduction coefficients should be insignificant, as the difference in difference estimate should not be significantly different between the two groups in the pre-treatment period (see Autor, 2003, for an application of this test). The event plots clearly indicate that all *before* coefficients are statistically insignificant and thereby support a common trend in the observations of both intervention and non-intervention EPAs, before the start of the new auction floor.

Ceilings to expansion

Another issue concerns the presence of ceilings to expand: if all land suitable for tobacco cultivation is exhausted, there are no possibilities for further growth of tobacco production. EPAs that meet these conditions cannot be used as controls. Potential availability of crop area is investigated by constructing estimates of extension of total crop area and area available by substitution of other crops (see Appendix E). Under fairly reasonable assumptions it is shown that most EPAs to have potential expansion opportunities for tobacco cultivation higher than 100% of existing tobacco area. The EPAs with less potential expansion opportunities still have a minimum opportunity for expanding tobacco area of 20%. Hence, average expansion

opportunities of crop area of non-intervention EPAs, expressed in terms of existing tobacco area, are no effective restriction.

Quality of the data

Researchers occasionally point at the poor quality of Malawi administrative data, mostly, however, in relation to maize production data (maize is occasionally referred to as a political crop). For this reason we have compared the EPA data from the Agro Economic Survey of the Ministry of Agriculture and Food Security (AES-MoAFS) – the data that we use for the empirical estimations – with the auction sales volume data from Tobacco Control Commission (TCC) and with tobacco information extracted from the Integrated Household Survey 2 (IHS-2) from the National Statistical Office (see Appendix B). Tobacco data from different sources clearly show discrepancies. Most discrepancies, however, have reasonable explanations¹². However, a number of discrepancies merit further investigation. Since the EPAs in the districts with the largest discrepancies (Dowa, Lilongwe, Kasungu and Mchinji) partly belong to the group of EPAs that is likely to benefit from the newly introduced auction floor, this observation points at the possibility of having estimated a statistical artefact in the impact estimations, that reflects the data collection process rather than a real response of tobacco growers. To investigate if the estimated impacts are statistical artefacts, we have checked the robustness of the results by omitting data from these districts.

The results reported in Appendix F show that estimates of impacts do change (which is not surprising given the key role these districts play in the identification of the impact of distance to auction floor) but remain to a large extent statistically significant and of similar size. These robustness checks confirm that previous estimation of impacts results can be maintained.

¹² Explanations are, for example: the distinction between smallholders and estates, and burley tobacco and other tobacco's, storage by farmers and lags in sales, measurement errors in recording, (illegal) cross border trade, etc.

6. Summary and conclusion

We have investigated the impact of improved market access for a typical developing country cash crop / export crop on the smallholder's decisions on cultivated area and production. For this purpose we have exploited tobacco area and production data for the period around the introduction of an additional tobacco auction floor in Malawi. Estimations are based on annual data by Extension Planning Area (EPAs), 198 in total, covering the whole of Malawi, for a period of seven years, from 2003 to 2009. Tobacco is the most important cash crop in Malawi, grown in all districts of Malawi, exclusively sold on auction floors, and subsequently entirely exported. There are four tobacco auction floors (Limbe (close to Blantyre), Kanengo (close to Lilongwe), Mzuzu and Chinkhoma), of which the one in Chinkhoma has started operations in 2004. The estimation results support a statistically significant positive impact of the introduction of the new auction floor and the related decrease in transport costs, on tobacco area and tobacco production. As the increase in production is larger than the increase in area, area productivity increases. The increase in area productivity suggests intensification of tobacco cultivation by existing growers. The impact of the introduction of the Chinkhoma auction floor is confirmed with generalised propensity matching, a matching technique that is especially designed for the case of continuous treatment. Alternative explanations for the estimated impact (estimated impact applies to all crops, common trend assumption, restrictions to expansion in non-intervention locations and the measured impact is the result of poor quality of the data) could all be rejected.

Acknowledgements

I would like to thank Eric Bartelsman, Hans Binswanger, Frank Bruinsma, Chris Elbers, Christopher Gilbert, Peter Lanjouw, Thea Nielsen, Menno Pradhan, Asger Moll Wingender, Bram Thuysbaert, and conference participants in G tenborg, Sweden (Nordic Conference in Development Economics, 2012), Dakar, Senegal (PEGNet, 2012), Universit  Dauphine, Paris, France (DIAL, 2013), Milan (ICAE, 2015), colleagues of VU-University Amsterdam and various anonymous journal reviewers for comments and suggestions on previous versions of this paper. Errors and omissions in the current paper are the responsibility of the author. I also would like to thank Hans Quene for skilful assistance in constructing the data and variables.

References

- Aker, J.C., 2010, 'Information for Markets Near and Far: Mobile Phones and Agricultural Markets in Niger', *American Economic Journal: Applied Economics*, 2 (July), 46-59.
- Balat, J., I.Brambilla, and G.Porto, 2009, 'Realizing the gains from trade: Export crops, marketing costs, and poverty', *Journal of International Economics*, 78, 1, 21-31.
- Bia, M. and A.Mattei, 2008, 'A Stata Package for the Estimation of the Dose-Response Function through adjustment for the generalized Propensity Score', *The Stata Journal*, 8, 3, 354-373.
- De Janvry, A., M.Fafchamps and E.Sadoulet, 1991, 'Peasant Household Behaviour with Missing Markets: Some Paradoxes Explained', *The Economic Journal*, 101, November, 1400-1417.
- De Janvry, A. and R.Kanbur (eds.), 2006, *Poverty, Inequality and Development, Essays in Honor of Erik Thorbecke*, Kluwer Publishing.
- De Janvry, A. and E.Sadoulet, 2006, 'Progress in the Modelling of Rural Households' Behaviour under Market Failures', Chapter 8 in De Janvry, A. and R.Kanbur, 2006.
- De Janvry, A. and E.Sadoulet, 2010, 'Agriculture for Development in Africa: Business-as Usual or New Departures', *Journal of African Economies*, 19, AERC, Sup. 2, ii7-ii39.
- Diagne, A. and M. Zeller, 2001, Access to Credit and its Impact on Welfare in Malawi, Research Report 116, IFPRI, Washington.
- Fafchamps, M., 1999, Rural poverty, Risk and Development, Rome, FAO.
- Fafchamps, M., and R.Vargas Hill, 2005, 'Selling at the Farm-gate or Travelling to the Market', *American Journal of Agricultural Economics*, 87, 3, August, 717-734.
- Food and Agricultural Organization of the United Nations (FAO), 2003, 'Issues in the Global Tobacco Economy: Selected case studies', Rome.
- Goetz, S.J., 1992, 'A Selectivity Model of Household Food Marketing Behavior in Sub-

- Saharan Africa', *American Journal of Agricultural Economics*, 74, 444-452.
- Goyal, A., 2010, 'Information, Direct Access to Farmers and Rural Market Performance in Central India', *American Economic Journal: Applied Economics*, 2 (July), 22-45.
- Harrigan, J., 2008, 'Food Insecurity, Poverty and the Malawian Starter Pack: Fresh Start or False Start?', *Food Policy*, 33, 237-249.
- Harrigan, J., 2003, 'U-Turns and Full Circles: Two Decades of Agricultural Reform in Malawi 1981-2000', *World Development*, 31, 5, 847-863.
- Hirano, K. and G. Imbens, 2004, 'The Propensity score with continuous treatment', in Gelman A. and Xiao-Li Meng (eds.), *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives*, John Wiley and Sons, Chapter 7, 73-84.
- Jacoby, H.G. and B. Minten, 2009, 'On measuring the Benefits of Lower Transport Costs', *Journal of Development Economics*, 89, 1, 28-38.
- Jaffee, S., 2003, 'Malawi's Tobacco Sector: Standing on One Strong Leg is Better Than on None', World Bank, Africa Region Working Paper Series, nr. 55.
- Jayne, T.S, 1994, 'Do High Food Marketing Costs Constrain Cash Crop Production? Evidence from Zimbabwe', *Economic Development and Cultural Change*, 42, 2 January, 387-402.
- Jensen, R., 2007, 'The Digital Divide: Information (Technology), Market Performance, and Welfare in the South Indian Fisheries Sector', *Quarterly Journal of Economics*, 72, 3, 879-924.
- Key, N., E.Sadoulet and A. de Janvry, 2000, 'Transactions Costs and Agricultural Household Supply Response', *American Journal of Agricultural Economics*, 82, May, 245-259.
- Kydd, J. and R.Christiansen, 1982, 'Structural Change in Malawi since Independence: Consequences of a Development Strategy Based on Large-Scale Agriculture', *World Development*, 10, 5, 355-375.

- Koester, U., G.Olney, C.Mataya, and T.Chidzanja, 2004, 'Status and Prospect of Malawi's Tobacco Industry: A Value Chain Analysis', report for The Emergency Drought Recovery Project, Ministry of Agriculture, Malawi.
- Lea, N. and L.Hammer, 2009, 'Constraints to Growth in Malawi', The World Bank, Policy Research Working Paper 5097.
- Negri, M and G.G.Porto, 2015, 'Burley Tobacco Clubs in Malawi: Nonmarket Institutions for Exports', *International Economics*, forthcoming.
- Minten, B. and S.Kyle, 1999, 'The effect of Distance and Road Quality on Food Collection, Marketing Margins, and Traders' Wages: Evidence from the Former Zaire', *Journal of Development Economics*, 60, 467-495.
- Omamo, S.W., 1998, 'Transport Costs and Smallholder Cropping Choices: An Application to Siaya District, Kenya', *American Journal of Agricultural Economics*, 80, 116-123.
- Orr, A., 2000, ' 'Green Gold'? : Burley Tobacco, Smallholder Agriculture, and Poverty Alleviation in Malawi', *World Development*, 28, 2, 347-363.
- Otañez, M.G., H.Mamudu and S.A.Glantz, 2007, 'Global Leaf Companies Control the Tobacco Market in Malawi', *Tobacco Control*, 16, 261-269.
- Poulton, C., J.Kydd and D.Kabame, 2007, 'All Africa Review of Experiences with Commercial Agriculture: Case Study on Malawi Tobacco', background paper for the Competitive Commercial Agriculture in Sub-Saharan Africa Study (CCAA study).
- Renkow, M., D.Hallstrom and D.Karanja, 2004, 'Rural Infrastructure, Transaction Costs and Market Participation in Kenya', *Journal of Development Economics*, 73, 349-367.
- Tchale, H. and J.Keyser, 2010, 'Quantitative Value Chain Analysis, An Application to Malawi', The World Bank, Policy Research Working Paper 5242.
- World Bank, 2004, Diagnostic Trade Integration Study, Volume 2, Malawi, Chapter 4, Tobacco Sector.

Appendix A Data used, data sources and variable construction

Annual data of smallholder agricultural production and crop area on the level of Extension Planning Area's (EPAs), for the years from 2003/04 to 2009/10, are from the Agro-Economic Survey of the Ministry of Agriculture and Food Security (AES-MoAFS). All production and area data pertain to smallholders and exclude estates. An EPA reclassification in Salima and Nkothakotha district has made a number of before-reclassification EPAs different from their equally named after-reclassification EPAs. Therefore, after reclassification observations – the shortest series – have been removed.

Monthly farm gate prices for tobacco are sourced from the Agro-Economic-Survey, Ministry of Agriculture and Food Security, and are observed for close to 50 locations, scattered over Malawi. However, these series are not complete. Around 43% of the tobacco farm gate price data used (annuals, seasonal averages) are directly taken from the data sources. The remaining observations are constructed by calculating a location specific (average) share of farm gate prices in national auction prices (Tobacco Control Commission (TCC)) and imputing these values to fill up the missing observations. Time series auction price data are unfortunately only available at the national level. Tobacco farm gate prices expressed as a share of the auction prices are 35.2% on average (median: 33.2%). This compares reasonably well with other sources (see Koester et al., 2004). All price data are attributed on the basis of proximity from markets to EPAs. In some cases this involved averaging over various locations (triangulation). Data on farm gate maize prices are also from the Agro-Economic-Survey. These monthly series are available for close to 60 locations, but in contrast with burley tobacco farm gate prices the maize price series are much more complete: around 84% of the maize farm gate price data are directly taken from the data source. The remaining observations – the missings – are constructed using nearby prices. Like tobacco prices the maize price series are attributed to EPAs on the basis of the minimum distance of the geographical location of farm

gate prices to the EPA. Groundnut prices are market prices – due to limited availability of farm gate prices and in contrast with tobacco and maize prices – and available for over 70 markets. All prices are deflated with the Malawi consumer price index for rural areas.

Annual data on rainfall in mm are from around 30 meteorological stations and supplied by the Department of Climate Change and Meteorological Services, Blantyre. Again we exploit the distance between meteorological stations and EPAs to find the rainfall series that is relevant for a specific EPA. The distance to the nearest weather station is, in most cases, less than 20km. If more than one weather station is nearest to an EPA we calculated the average between the nearest weather stations (triangulation).

Data on the number of households by EPA for one year (2007/08) are from the Ministry of Agriculture and Food Security. Combined with district data on average household size and population, we have constructed EPA population series for 2003/04 to 2009/10. Population by district data are census based and from the National Statistical Office (NSO). The EPA population series is used to construct EPA population density (EPA population in numbers by EPA land area in km²) or, alternatively, per capita area. EPA land area is constructed on the basis of a map of EPAs and made consistent with data on district area (source: www.geohive.com). The size of EPAs in km², pertains to land area (and hence excludes large lakes, like for example Lake Chilwa). The population density series varies both over time and between EPAs (but, naturally, the variation over time is limited). For the construction of an agglomeration index and distance to cities and towns we use a 1998 and 2008 listing of population of Malawi cities and towns, taken from National Statistical Office of Malawi.

For spatial integration of EPAs with the rest of the country we have constructed an agglomeration index. The agglomeration index is population size of city/town over distance to city/town, summed over all cities and towns, or in formula:

$$A_{ij} = \frac{1}{Y} [L_i K_j R_j / @P_A R_j ?B N R Y]$$

SDA EN A , F = B N ?B =@ P = AN

The larger the population of the city or town and the shorter the distance to the city or town, the higher the index. The agglomeration index reflects the degree of embedding of an EPA in the network of cities and towns. The higher this degree the lower transaction costs, and hence, we expect a positive relationship with tobacco production per hectare, production and area. It also represents the requirement to have access to market, since cash crop growers need to rely more on markets for their food purchases, for purchases of manufactured goods and for purchases of inputs (see de Janvry and Sadoulet, 2006)).

For descriptive purposes we use one year of auction transaction data (2009, a total of around 62,000 transactions), which were kindly made available by the Tobacco Control Commission (TCC). Each transaction contains information on type of tobacco, number of bales, volume in kg, value in US\$, district of origin and club. Also for descriptive purposes we use annual aggregate times series data on the tobacco sector from TCC, posted on TCC website (www.tccmw.com).

Appendix B Comparison of tobacco data by source: AES-MoAFS, IHS (NSO) and TCC*

Figure B1 Aggregate annual production (AES-MoAFS) versus sales volume (TCC)

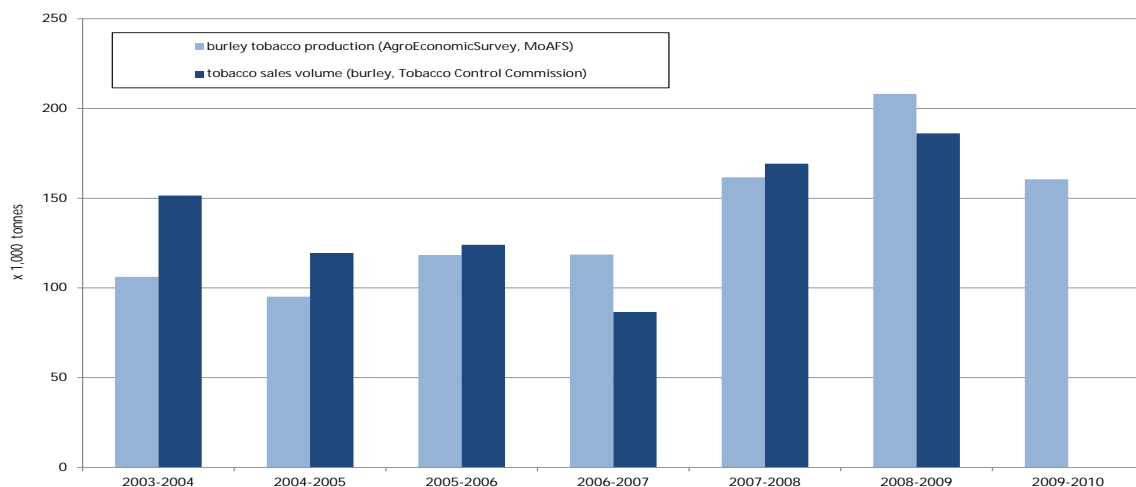


Figure B2 Production (AES-MoAFS) versus sales volume (TCC) by district, 2009

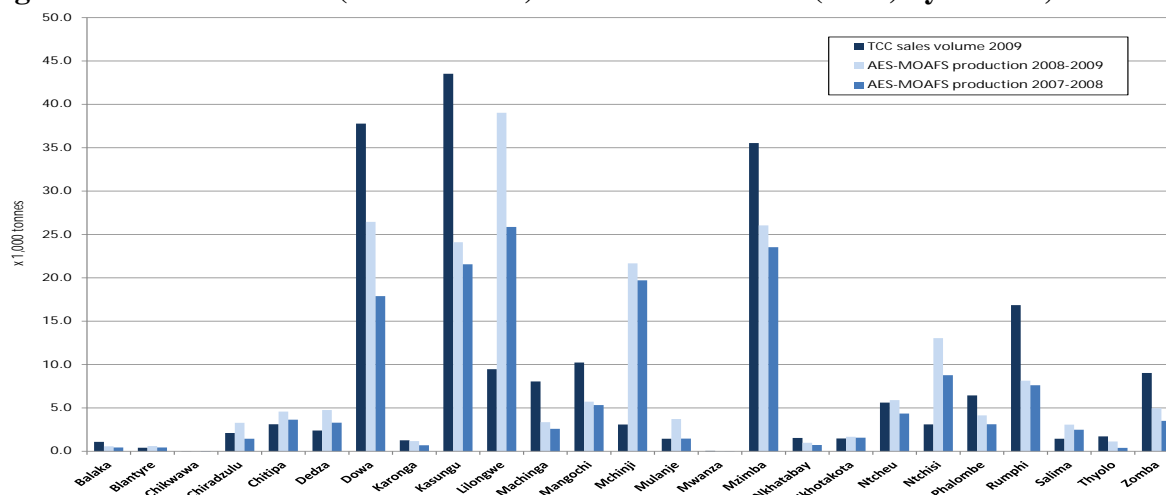
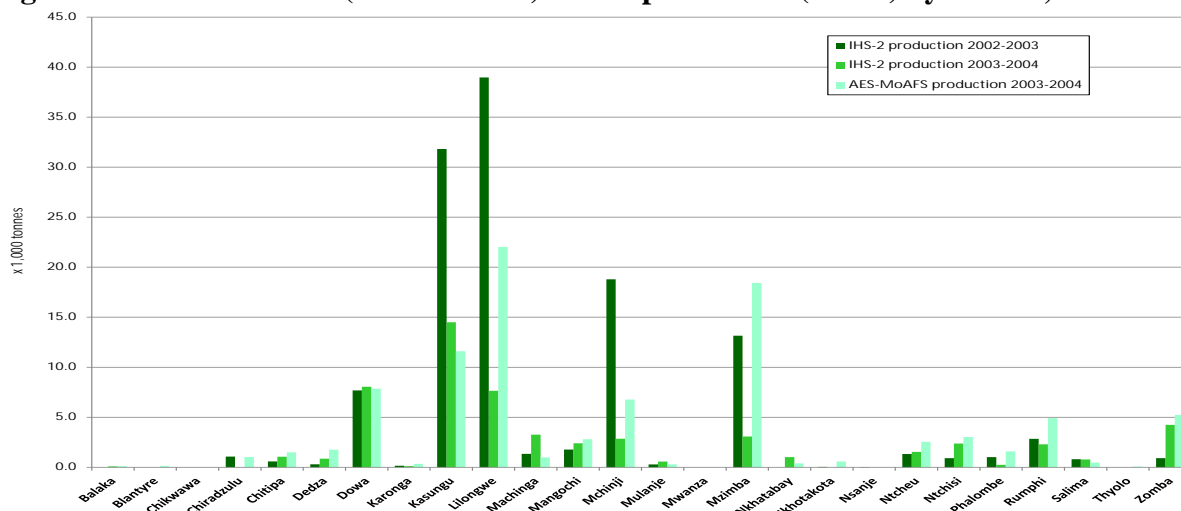


Figure B3 Production (AES-MoAFS) versus production (IHS-2) by district, 2003/04

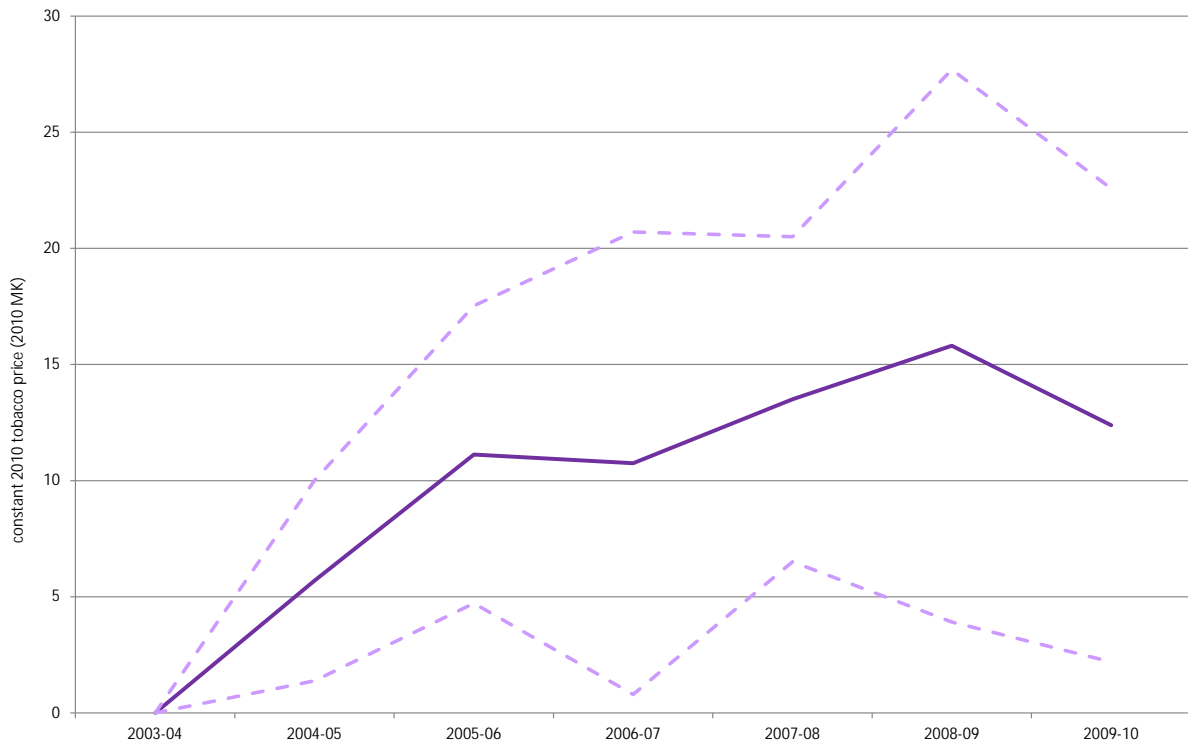


* AES-MoAFS: Agro-Economic Survey, Ministry of Agriculture and Food Security; TCC: Tobacco Control Commission; IHS: Integrated Household Survey; NSO: National Statistical Office.

Appendix C Event study plot for farm gate prices

The Figure below shows the plot of the event estimation of tobacco farm gate prices and is similar to the plots in the main text (Figure 3). Likewise, we estimated $\hat{U}_{it} = \alpha_0 + \alpha_1 \Delta t + \alpha_2 \Delta t^2 + \alpha_3 \Delta t^3 + \alpha_4 \Delta t^4 + \alpha_5 \Delta t^5 + \alpha_6 \Delta t^6 + \alpha_7 \Delta t^7 + \alpha_8 \Delta t^8 + \alpha_9 \Delta t^9 + \alpha_{10} \Delta t^{10} + \alpha_{11} \Delta t^{11} + \alpha_{12} \Delta t^{12} + \alpha_{13} \Delta t^{13} + \alpha_{14} \Delta t^{14} + \alpha_{15} \Delta t^{15} + \alpha_{16} \Delta t^{16} + \alpha_{17} \Delta t^{17} + \alpha_{18} \Delta t^{18} + \alpha_{19} \Delta t^{19} + \alpha_{20} \Delta t^{20} + \alpha_{21} \Delta t^{21} + \alpha_{22} \Delta t^{22} + \alpha_{23} \Delta t^{23} + \alpha_{24} \Delta t^{24} + \alpha_{25} \Delta t^{25} + \alpha_{26} \Delta t^{26} + \alpha_{27} \Delta t^{27} + \alpha_{28} \Delta t^{28} + \alpha_{29} \Delta t^{29} + \alpha_{30} \Delta t^{30} + \alpha_{31} \Delta t^{31} + \alpha_{32} \Delta t^{32} + \alpha_{33} \Delta t^{33} + \alpha_{34} \Delta t^{34} + \alpha_{35} \Delta t^{35} + \alpha_{36} \Delta t^{36} + \alpha_{37} \Delta t^{37} + \alpha_{38} \Delta t^{38} + \alpha_{39} \Delta t^{39} + \alpha_{40} \Delta t^{40} + \alpha_{41} \Delta t^{41} + \alpha_{42} \Delta t^{42} + \alpha_{43} \Delta t^{43} + \alpha_{44} \Delta t^{44} + \alpha_{45} \Delta t^{45} + \alpha_{46} \Delta t^{46} + \alpha_{47} \Delta t^{47} + \alpha_{48} \Delta t^{48} + \alpha_{49} \Delta t^{49} + \alpha_{50} \Delta t^{50} + \alpha_{51} \Delta t^{51} + \alpha_{52} \Delta t^{52} + \alpha_{53} \Delta t^{53} + \alpha_{54} \Delta t^{54} + \alpha_{55} \Delta t^{55} + \alpha_{56} \Delta t^{56} + \alpha_{57} \Delta t^{57} + \alpha_{58} \Delta t^{58} + \alpha_{59} \Delta t^{59} + \alpha_{60} \Delta t^{60} + \alpha_{61} \Delta t^{61} + \alpha_{62} \Delta t^{62} + \alpha_{63} \Delta t^{63} + \alpha_{64} \Delta t^{64} + \alpha_{65} \Delta t^{65} + \alpha_{66} \Delta t^{66} + \alpha_{67} \Delta t^{67} + \alpha_{68} \Delta t^{68} + \alpha_{69} \Delta t^{69} + \alpha_{70} \Delta t^{70} + \alpha_{71} \Delta t^{71} + \alpha_{72} \Delta t^{72} + \alpha_{73} \Delta t^{73} + \alpha_{74} \Delta t^{74} + \alpha_{75} \Delta t^{75} + \alpha_{76} \Delta t^{76} + \alpha_{77} \Delta t^{77} + \alpha_{78} \Delta t^{78} + \alpha_{79} \Delta t^{79} + \alpha_{80} \Delta t^{80} + \alpha_{81} \Delta t^{81} + \alpha_{82} \Delta t^{82} + \alpha_{83} \Delta t^{83} + \alpha_{84} \Delta t^{84} + \alpha_{85} \Delta t^{85} + \alpha_{86} \Delta t^{86} + \alpha_{87} \Delta t^{87} + \alpha_{88} \Delta t^{88} + \alpha_{89} \Delta t^{89} + \alpha_{90} \Delta t^{90} + \alpha_{91} \Delta t^{91} + \alpha_{92} \Delta t^{92} + \alpha_{93} \Delta t^{93} + \alpha_{94} \Delta t^{94} + \alpha_{95} \Delta t^{95} + \alpha_{96} \Delta t^{96} + \alpha_{97} \Delta t^{97} + \alpha_{98} \Delta t^{98} + \alpha_{99} \Delta t^{99} + \alpha_{100} \Delta t^{100} + \epsilon_{it}$ and plotted \hat{U}_{it} over t years. We are hesitant to put much trust in the plotted event figure, because data on tobacco farm gate prices are to large extent constructed (see Appendix A). But, apart from the wide spread, the plot is quite supportive for our investigations: It confirms statistically significant average higher tobacco farm gate prices due to the introduction of a new auction floor, and corresponds with around 20% higher farm gate prices.

Figure C1 Impact of lower transport costs on farm gate tobacco prices



Appendix D Finding evidence for substitution and expansion

Is the increase in tobacco area realized by substitution of area of other crops, or the result of area expansion? In order to investigate this we propose the following decomposition: Total crop area is the sum of tobacco area and non-tobacco area ($a_{all,t} = a_{tb,t} + a_{ntb,t}$) where non-tobacco area is an aggregate of all alternative crops including maize, groundnuts and pulses.

The share of tobacco area in total crop area, hence, $a_{tb,t} = a_{all,t} - (1 - \alpha) a_{all,t}$.

After lagging and differencing, we write the change in tobacco area as: $\hat{a}_{b,t} = a_{all,t} - a_{all,t-1} -$

$(1 - \alpha) a_{all,t} - (1 - \alpha) a_{all,t-1}$, I W K H $(1 - \alpha)$ and all changes in tobacco area stem from a change in total crop area, then $\hat{a}_{b,t} = \alpha (a_{all,t} - a_{all,t-1})$. If total crop area is constant ($a_{all,t} = a_{all,t-1}$) and all changes in tobacco area stem from a change in the share of tobacco, then $\hat{a}_{b,t} = (1 - \alpha) a_{all,t}$. On the basis of this decomposition we regress the change in tobacco area on the share of tobacco area (α) and total crop area ($a_{all,t}$) to find the average contribution of expansion and substitution as the coefficients of these regressors.

The estimation results for the intervention locations during the intervention years (column 4) indicate that the change in tobacco area correlates positively and statistically significant with the previous year share of tobacco area. Correlation with previous year total crop area is at most weakly significant (if we include EPA fixed effects). Hence, both estimations indicate that variation in tobacco area is dominated by variation in total crop area. The estimations do not support substitution of tobacco area with area of other crops.

Table D1 Substitution and expansion of tobacco crop area

dependent variable	F	K	D	Q	J	H	(L)	Q	W	R	E	D	F
Explanatory variables:	(1)	(2)	(3)	(4)									
total crop area ($a_{\text{all crops},t}$)	0.0256*	0.0638***					0.0030					0.0290*	
	(0.0138)	(0.0196)					(0.0092)					(0.0158)	
tobacco area share ($a_{\text{tb},t}/a_{\text{all crops},t}$)	12238***	17024**					12291***					24957***	
	(2955.0)	(6255.4)					(1209.7)					(4346.9)	
dEPA _i	yes	yes					yes					yes	
crop years	2003/04 - 2005/06	2003/04 - 2005/06					2006/07 - 2009/10					2006/07 - 2009/10	
(non) intervention EPAs	non-intervention	intervention					non-intervention					intervention	
number of observations	294	77					592					157	
R ²	0.782	0.700					0.479					0.525	

Notes – The table reports estimates total crop area and the share of tobacco area in total crop area, on tobacco crop area. Estimations are based on annual data from 2003-04 to 2009-10. All estimations include EPA and year fixed effects. All equations are estimated with OLS. Robust standard errors clustered by EPAs are given in parentheses (.) below the coefficient. R² = coefficient of determination. * means significant at the 10% level ($p < 0.10$), ** at the 5% level ($p < 0.05$), *** at the 1% level ($p < 0.01$).

Appendix E Is crop area available for tobacco cultivation effectively limited?

Potential availability of crop area can be investigated: EPA data on crop area are available for all major crops and for this exercise we distinguish crop area for tobacco, maize and other crops, where “other crops” is an aggregate of rice, groundnuts, pulses, cassava, sweet potatoes, cotton, sorghum and millet. The data underscore the overwhelming importance of maize, with an average share of around 50% of total crop area. Note that the average tobacco share in total crop area is less than 5% (median less than 2.5%). Expansion of tobacco crop area occurs either through expansion of total crop area or through substitution with other crops. We assume that the potential for expansion of total crop area may be approximated with the gap between the maximum total crop area realized over the years 2003/04 to 2009/10 and actual total crop area, both by EPA. Hence, $L_t^{\text{pot}} = \max_{i \in \{2003/04, \dots, 2009/10\}} A_{i,t} - A_{t,t}$ where L_t^{pot} is potential crop area by expansion and $A_{t,t}$ is crop area of all crops, both in EPA i and in period t . Next, we assume that all crop area potentially available through substitution with crop area of other crops originates from the “other crops” sector and only to the extent that crop area allocated to other crops exceeds a minimum level, reflecting minimum requirements for food security. By restricting area available for potential substitution to a limited part of the “other crops” area, we acknowledge the importance of maize and other food crops on food security grounds. Hence, crop area potentially available through substitution with other crops is computed as the difference between current “other crop” area and the minimum area allocated to other crops during the period 2003/04 to 2009/10. Hence, we have $L_t^{\text{sub}} = \max\{0, A_{t,t} - \min_{i \in \{2003/04, \dots, 2009/10\}} A_{i,t}^{\text{other}}\}$ where L_t^{sub} is potential crop area by substitution and $A_{t,t}^{\text{other}}$ is crop area allocated to other crops, both in EPA i and in period t . The sum of potential expansion and substitution area ($L_t^{\text{pot}} + L_t^{\text{sub}}$) expressed in terms of tobacco area ($L_t^{\text{pot}} + L_t^{\text{sub}}$)

NAZ (t), should be high in order not to be a restriction for growth of tobacco area. On the basis of this numerical exercise we find a few of the control EPAs to have potential expansion opportunities for tobacco cultivation less than 100% of existing tobacco area in the period 2005/06 to 2009/10 (EPAs in the districts Lilongwe (3), Machinga (1), Mzimba (1) and Rumphi (1). Even these EPAs have a minimum opportunity for expanding tobacco area of 20%. Hence, the average expansion opportunities of non-intervention EPAs, expressed in terms of existing tobacco area, are high and we should conclude that there are no effective restrictions in this respect¹³.

¹³ Strictly we should also analyse if availability of labour and tobacco cultivation expertise is a restriction to growth of tobacco production in the control EPAs. This is particularly relevant since tobacco cultivation is considered labour intensive. Unfortunately we are unable to implement such an analysis. Hence, we have implicitly assumed no restrictions on these grounds.

Appendix F Investigating the quality of the data

Table D1 Market access in tobacco: omitting observations of specific district

Dependent variable:	ln(production)	ln(area)	ln(production per hectare)
District omitted: Kasungu	(2)	(3)	(1)
ln(distance to auction floor)	-0.937*** (0.340)	-0.376* (0.226)	-0.561* (0.288)
Number of observations	1144	1144	1144
R ²	0.939	0.958	0.510
District omitted: Lilongwe	(2)	(3)	(1)
ln(distance to auction floor)	-0.889** (0.360)	-0.340 (0.236)	-0.550* (0.308)
Number of observations	1026	1026	1026
R ²	0.937	0.956	0.514
District omitted: Mzimba	(2)	(3)	(1)
ln(distance to auction floor)	-0.920** (0.465)	-0.085 (0.207)	-0.835* (0.448)
Number of observations	1033	1033	1033
R ²	0.941	0.961	0.517
District omitted: Mchinji	(2)	(3)	(1)
ln(distance to auction floor)	-0.894*** (0.293)	-0.485** (0.235)	-0.409** (0.178)
Number of observations	1117	1117	1117
R ²	0.938	0.456	0.515

Notes – The table reports estimates of ATE of distance to auction floor on tobacco production, crop area and production per hectare. The robustness of the basic estimation of Table 2 is tested by omitting observations from specific districts. Estimations are based on annual data from 2003-04 to 2009-10 (seven years). All estimations include EPA and year fixed effects. All equations are estimated with OLS. Robust standard errors clustered by EPAs are given in parentheses (.) below the coefficient. R² = coefficient of determination.

Appendix G Tobacco auction transactions

Table G1 Sales volume by tobacco auction floor* and by district of origin, 2009

District	Sales volume (in 1000 kg)					per district distribution over auctions in %					per auction distribution over districts in %				
	MZZ	CNK	LIL	LMB	TOTAL	MZZ	CNK	LIL	LMB	TOTAL	MZZ	CNK	LIL	LMB	TOTAL
Chitipa	2991	0.1	118	0.1	3110	96.2	0.0	3.8	0.0	100	4.3	0.0	0.2	0.0	1.5
Karonga	12288	16	24	4.0	1272	96.5	1.3	1.9	0.3	100	1.8	0.1	0.0	0.0	0.6
Rumphi	16383	469		4.0	16856	97.2	2.8		0.0	100	23.6	1.7		0.0	8.2
Nkhatabay	1530	3		2.7	1535	99.6	0.2		0.2	100	2.2	0.0		0.0	0.7
Mzimba	34913	550	34	46	35543	98.2	1.5	0.1	0.1	100	50.2	2.0	0.1	0.1	17.2
Nkhotakota	472	568	29	403	1472	32.1	38.6	1.9	27.4	100	0.7	2.0	0.1	0.7	0.7
Kasungu	9228	16737	16424	1139	43528	21.2	38.5	37.7	2.6	100	13.3	60.2	30.1	2.1	21.1
Ntchisi	415	1977	73	633	3098	13.4	63.8	2.4	20.4	100	0.6	7.1	0.1	1.2	1.5
Dowa	1361	5494	29337	1593	37786	3.6	14.5	77.6	4.2	100	2.0	19.8	53.7	2.9	18.3
Mchinji	400	678	421	1580	3079	13.0	22.0	13.7	51.3	100	0.6	2.4	0.8	2.9	1.5
Salima	54	45	4.9	1345	1450	3.7	3.1	0.3	92.8	100	0.1	0.2	0.0	2.5	0.7
Lilongwe	481	1199	6364	1428	9471	5.1	12.7	67.2	15.1	100	0.7	4.3	11.7	2.6	4.6
Dedza	4.1	22	1733	629	2388	0.2	0.9	72.6	26.3	100	0.0	0.1	3.2	1.1	1.2
Ntcheu	7.4	4.7	13	5583	5608	0.1	0.1	0.2	99.5	100	0.0	0.0	0.0	10.2	2.7
Mangochi	4.8	1.8		10225	10232	0.0	0.0		99.9	100	0.0	0.0		18.6	4.9
Machinga	2.4	1.5	2.2	8050	8056	0.0	0.0	0.0	99.9	100	0.0	0.0	0.0	14.7	3.9
Balaka		1.6	0.3	1080	1082		0.1	0.0	99.8	100		0.0	0.0	2.0	0.5
Zomba	10	2.1		9015	9027	0.1	0.0		99.9	100	0.0	0.0		16.4	4.4
Mwanza				48	48				100.0	100				0.1	0.0
Blantyre	5.9	3.0	3.3	390	402	1.5	0.8	0.8	96.9	100	0.0	0.0	0.0	0.7	0.2
Chiradzulu	13	11	6.7	2071	2102	0.6	0.5	0.3	98.6	100	0.0	0.0	0.0	3.8	1.0
Phalombe	0.4	3.2		6429	6433	0.0	0.0		99.9	100	0.0	0.0		11.7	3.1
Mulanje	0.7			1441	1442	0.1			99.9	100	0.0			2.6	0.7
Thyolo			0.0	1714	1714			0.0	100.0	100			0.0	3.1	0.8
Chikwawa	0.0	0.0	0.0	19	19	0.0	0.0	0.0	100.0	100				0.0	0.0
Total	69504	27788	54588	54871	206750	33.6	13.4	26.4	26.5	100	100	100	100	100	100

Source: transaction data for 2009 from the Tobacco Control Commission, Malawi (around 60,000 observations); * CNK = Chinkhoma; LIL = Lilongwe (Kanengo); LMB = Limbe; MZZ = Mzuzu; Notes – Districts of origin and auction floors, ordered from north to south. Regions (north, central and south) are distinguished by shading.